

# Unsaturated subsurface flow with surface water and nonlinear in- and outflow conditions\*

Heiko Berninger<sup>†</sup> Mario Ohlberger<sup>‡</sup> Oliver Sander<sup>§</sup> Kathrin Smetana<sup>¶</sup>

January 14, 2013

## Abstract

We analytically and numerically analyze groundwater flow in a homogeneous soil described by the Richards equation, coupled to surface water represented by a set of ordinary differential equations (ODE's) on parts of the domain boundary, and with nonlinear outflow conditions of Signorini's type. The coupling of the partial differential equation (PDE) and the ODE's is given by nonlinear Robin boundary conditions. This article provides two major new contributions regarding these infiltration conditions. First, an existence result for the continuous coupled problem is established with the help of a regularization technique. Second, we analyze and validate a solver-friendly discretization of the coupled problem based on an implicit-explicit time discretization and on finite elements in space. The discretized PDE leads to convex spatial minimization problems which can be solved efficiently by monotone multigrid. Numerical experiments are provided using the DUNE numerics framework.

**Keywords:** saturated-unsaturated porous media flow, Kirchhoff transformation, convex minimization, finite elements, monotone multigrid, nonlinear transmission problem

**AMS Subject Classification:** 35K61, 65N30, 65N55, 76S05

---

\*This work was supported by the BMBF under contract numbers 03MOPAF1 and 03KOPAF4

<sup>†</sup>Section de Mathématiques, Université de Genève, 2-4 rue du Lièvre, CP 64, 1211 Genève 4, Switzerland, Heiko.Berninger@unige.ch

<sup>‡</sup>Institute for Computational and Applied Mathematics, University of Muenster, Einsteinstr. 62, 48149 Muenster, Germany, mario.ohlberger@uni-muenster.de

<sup>§</sup>Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany, sander@igpm.rwth-aachen.de

<sup>¶</sup>Institute for Computational and Applied Mathematics, University of Muenster, Einsteinstr. 62, 48149 Muenster, Germany, kathrin.smetana@uni-muenster.de

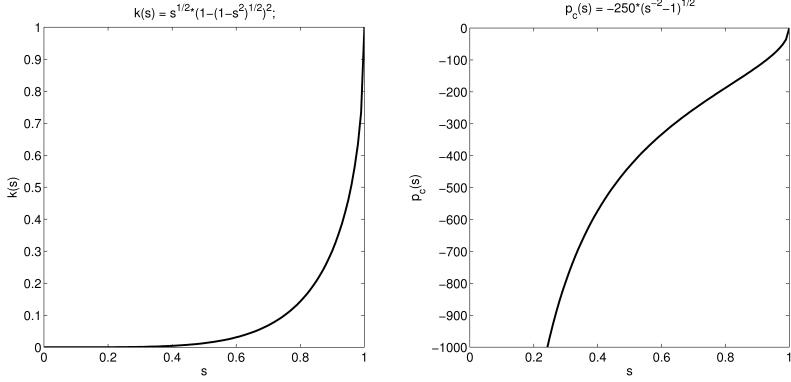


Figure 1: Typical shapes of the coefficient functions  $k(s)$  and  $p_c(s)$

## 1 Introduction

This article is concerned with existence of solutions, and efficient numerical approximation of unsaturated groundwater flow with surface water and nonlinear in- and outflow conditions. To start with, let us introduce the considered model. Let  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$  denote a bounded domain occupied by a homogeneous soil with boundary  $\partial\Omega$  of class  $C^1$  and outer normal  $\nu$ . Let  $\Sigma_{\text{in}}$ ,  $\Sigma_{\text{out}}$  and  $\Sigma_N \subset \partial\Omega$  denote three relatively open, pairwise disjoint  $d - 1$ -dimensional  $C^1$ -manifolds, representing the infiltration, outflow and Neumann boundaries, such that  $\partial\Omega = \overline{\Sigma_{\text{in}}} \cup \overline{\Sigma_{\text{out}}} \cup \overline{\Sigma_N}$  and  $\Sigma_{\text{in}} \cap \Sigma_{\text{out}} = \emptyset$ . The unsaturated groundwater flow is modelled in the time interval  $[0, T]$  using Richards equation (see e.g. [9]) for the water saturation  $s : \Omega \times [0, T] \rightarrow [0, 1]$  and the water pressure  $p : \Omega \times [0, T] \rightarrow \mathbb{R}$

$$n\partial_t s - \operatorname{div}(k(s)\mu^{-1}(\nabla p + \mathbf{e})) = f \quad \text{in } \Omega \times [0, T]. \quad (1)$$

Here  $n$  denotes the porosity,  $\mu$  the viscosity,  $k$  the permeability, and  $\mathbf{e}$  the gravity vector. For simplicity we set  $n = \mu = 1$  and  $\mathbf{e} = (0, 0, 1)^T$  in this paper. Due to our assumption of a homogeneous soil, the permeability  $k$  depends only on the saturation  $s$ . In the Richards model, the air pressure in the pore space is assumed to be constant. We consider it normalized to  $p_{\text{gas}} = 0$  and replace the water pressure in (1) by the capillary pressure function  $\tilde{p}_c(s)$ . Thus, the capillary pressure  $p_c$  has to be seen as a multivalued graph with  $p_c(s) = \{\tilde{p}_c(s)\}$  for  $s < 1$  and  $p_c(1) = [\tilde{p}_c(1), \infty)$ . Furthermore, we assume the residual saturation, meaning the saturation level below which no flow of water occurs, to be zero, such that  $p_c$  is defined on  $[0, 1]$ . This is necessary for the definition of the infiltration boundary conditions. Additionally, we suppose  $\tilde{p}_c(1) = 0$ , which is essential for the limiting process on the outflow boundary. Typical shapes of  $k(s)$  and  $p_c(s)$  are depicted in Fig. 1. It can be seen that the coefficient functions degenerate in the sense that  $k(0) = 0$  and that  $k$  and  $p_c$  may have slopes that are unbounded in the neighbourhood of  $s = 1$  and  $s = 0$ , respectively. These degeneracies of the coefficient functions are one of the major challenges when treating (1) analytically.

We impose three different types of boundary conditions. On  $\Sigma_N$  we prescribe homogeneous Neumann boundary conditions. On the outflow boundary  $\Sigma_{\text{out}}$  we

choose the Signorini-type boundary conditions

$$\mathbf{q} \cdot \boldsymbol{\nu} \geq 0, \quad p \leq 0, \quad (\mathbf{q} \cdot \boldsymbol{\nu}) p = 0 \quad \text{on } \Sigma_{\text{out}} \times [0, T] \quad (2)$$

with  $\mathbf{q} = -k(s)(\nabla p + \mathbf{e})$  the Darcy velocity of the fluid. This models a hydrophilic porous medium that is in contact with air (cf. [31]). The inequalities (2) express that (i) water may not enter the domain, (ii) the pressure cannot be positive as  $p_{\text{air}}$  has been normalized to 0, and (iii) water can exit only if  $p = 0$ . Concerning the analysis given below, a major problem is to give sense to the product of traces in the last equality of (2). On the infiltration boundary  $\Sigma_{\text{in}}$ , we only consider inflow of water into the soil and ponding of water on the surface (without runoff). We assume that there is an external water source  $r(x, t)$ ,  $x \in \Sigma_{\text{in}}$ . The water at  $x$  either infiltrates directly into the soil, or accumulates in  $x$  (ponding). Hence the state of the surface water can be described by the water table  $w : \Sigma_{\text{in}} \times [0, T] \rightarrow \mathbb{R}$ . At the absence of surface water runoff,  $w$  is modelled as an ensemble of ordinary differential equations

$$\partial_t w = \mathbf{q} \cdot \boldsymbol{\nu} + r \quad \text{on } \Sigma_{\text{in}} \times [0, T], \quad (3)$$

which is coupled to the subsurface boundary flux  $\mathbf{q}$  (to be understood in a suitable weak sense). Note that although  $\Sigma_{\text{in}}$  is the only part of the boundary through which inflow can occur, we may also have outflow through  $\Sigma_{\text{in}}$ . A possible model for the flux  $\mathbf{q}$  on  $\Sigma_{\text{in}}$  is to consider it as a vertical leakage through a very small semipervious layer of thickness  $b$  and hydraulic conductivity  $K_h$  [24], which can be modelled (cf. [9, Chapt. 6.4]) as

$$\mathbf{q} \cdot \boldsymbol{\nu} = \frac{p - w}{c}. \quad (4)$$

Here  $c = b/K_h$  is the resistance of the semipervious layer. Its inverse  $c^{-1}$  is also known as leakage coefficient. Unfortunately, (4) is only valid for  $w > 0$ , as in the case of  $w = 0$  and a negative pressure  $p$ , we would obtain an inflow into the ground without water present on the infiltration boundary. Following [24] we modify (4) by a  $w$ -dependent factor to obtain a new law valid for all  $w$

$$\mathbf{q} \cdot \boldsymbol{\nu} = \frac{p_+ + p_- \min\{1, (\frac{w}{\sigma})_+\} - w}{c}, \quad (5)$$

where  $p_+ = \max\{p, 0\}$  and  $p_- = \min\{p, 0\}$ . Now,  $w = 0$  in (5) implies  $\mathbf{q} \cdot \boldsymbol{\nu} = 0$  if  $p \leq 0$ . The parameter  $\sigma$  is a regularization threshold, with (5) being the same as (4) for  $w \geq \sigma$ . As we will see later, solutions of this model fulfill  $w \geq 0$  automatically.

Existence results for unsaturated flow in porous media go back to the pioneering work of Alt, Luckhaus, and Visintin [2]. The basis of their approach - which we will also use in this article - is the so called Kirchhoff transformation. [37] picks up the methods used in [2] and proves the existence of solutions for (1) by regularization techniques, with a stronger solution concept than in [2]. The vital improvement is that [37] allows dry regions, meaning regions where the saturation lies below the (positive) residual saturation, on the outflow boundary. Hence, one has to modify the boundary conditions, as the pressure is not defined in dry regions. In [37] Robin boundary conditions are used in the regularized problem and convergence to the limit on the outflow boundary is obtained via

defect measures. Without considering outflow conditions, [18] also studies dry regions and proves the continuity of the free boundaries between the saturated and unsaturated regions and between the unsaturated and dry regions. In [33] uniqueness of the solution is proved for the problem introduced in [2] by adapting the methods developed in [32]. In order to show  $L^1$ -contraction and with this the uniqueness of the solution, doubling of variables techniques are used as introduced in [29]. With respect to the condition on the infiltration boundary, in [24] condition (5) is used to couple the Richards equation with a hyperbolic PDE modelling ponding of water and surface runoff on the infiltration boundary. Under the restriction  $\partial_t s \in L^1(\Omega)$  existence and uniqueness is shown for this coupled system. We remark that we do not need this assumption in our approach.

Concerning the numerical treatment of the Richards equation, a rich literature can be found. For an overview we refer to [10, Sec. 2.2] and [13] and the literature cited therein. The Richards equation has been discretized using finite volume methods [22, 27], mixed finite element methods [3, 35], finite elements [25, 26], and discontinuous Galerkin schemes [5, 20]. In most cases, the resulting algebraic systems were solved using Newton's method, which, however, suffers from ill-conditioning problems due to the degeneracies in the parameter functions (Fig. 1). In contrast, Berninger et al. [13] discretize the Kirchhoff-transformed Richards equation in a way that the resulting spatial problems can be solved efficiently using a monotone multigrid method. Since this approach is based on convexity rather than on smoothness, it is also well-suited for non-smooth Signorini-type boundary conditions. Numerical studies in [13] demonstrate the efficiency of the solver, as well as robustness for extreme soil parameters which correspond to parameter functions that degenerate into step functions. An extension of this approach to heterogeneous soil using domain decomposition techniques can be found in [12]. Concerning the numerical coupling of Richards equation with surface water, we mention [20, 21, 39], where shallow water equations are used as the surface water model. Although different discretizations are applied, all these approaches enforce continuity of pressure and normal flux, i.e., mass conservation, across the interface. These coupling conditions have also been considered in an approach based on the Kirchhoff transformed subsurface flow in [8] and [14], where an implicit time discretization of the coupled subsurface and surface water models has been solved with a heterogeneous domain decomposition method. In [14], however, the pressure continuity was replaced by a leakage condition such as (4) that represented clogging of a nearly impermeable river bed. Accordingly, a Robin–Neumann type iteration was applied in [14], whereas in [8] a Dirichlet–Neumann type iteration was used to solve the coupled system. In the present article, we choose an explicit time discretization for the surface water as in [10, Sec. 4.3].

Let us now give an outline of the article. In the first part of this article we prove a new existence result for the Richards equation (1) endowed with nonlinear outflow conditions of Signorini's type (2) and coupled to surface water by nonlinear Robin conditions (3, 5). The proof is done without requiring the assumption  $\partial_t s \in L^1(\Omega)$ . To achieve this goal, in Section 2 we introduce a global pressure  $u$  with the help of the Kirchhoff transformation and reformulate the system accordingly. In Section 3, we state the main existence theorem and derive its proof by combining two central approaches for existence and uniqueness results for nonlinear, degenerated parabolic PDEs, namely the regularization

technique (cf. [31, 37]) and the method of  $L^1$ -contraction (cf. [24, 33]).

In the second part of the article, starting with Section 4, we present our discretization of the coupled nonlinear problem that is implicit-explicit in time and uses finite elements in space. Via variational inequalities, the implicit-explicit time discretization leads to spatial convex minimization problems which contain nonlinear outflow and nonlinear physical Robin boundary conditions but no Dirichlet conditions. We prove unique solvability of these problems if the relative permeability is non-degenerate (e.g., regularized) and the water table is non-zero. In Section 5, we give a numerical example combining the coupled surface-subsurface problem and Signorini outflow conditions. We observe that the proposed time discretization leads to a stable method for reasonable time step sizes. We also observe that the surface water height remains nonnegative as predicted by the theory, with the exception of certain singular points. Our numerical approach proves to be useful even beyond the confines of the assumptions needed for the existence result.

## 2 Global pressure formulation of the fully coupled system and assumptions on the data

For a better assessment of the unsaturated flow equation, we use Kirchoff's transformation to obtain an equivalent formulation for a global pressure  $u$ . We then summarize the resulting fully coupled system with nonlinear boundary conditions and detail all assumptions on the data that we will need for our analysis.

Let us define the Kirchoff transformation

$$\tilde{\Phi}(s) := \int_0^{p_c(s)} k(p_c^{-1}(q)) dq, \quad \Phi(s) = \begin{cases} \tilde{\Phi}(s) & \text{if } s < 1, \\ [\tilde{\Phi}(1), \infty) & \text{if } s = 1, \end{cases} \quad (6)$$

which yields  $\nabla \Phi(s) = k(s) \nabla p_c(s)$ . We then define the global pressure function

$$u(x, t) := \begin{cases} \tilde{\Phi}(s(x, t)) & \text{if } s(x, t) < 1, \\ \tilde{\Phi}(1) + k(1)(p(x, t) - \tilde{p}_c(1)) & \text{if } s(x, t) = 1. \end{cases}$$

We thus have  $\nabla u = k(s) \nabla p$  for a sufficiently smooth solution  $(s, p)$  of (1). Therefore, we obtain as equivalent versions of the Richards equation

$$\begin{aligned} n\partial_t s &= \mu^{-1} \operatorname{div}(\nabla u + k(s) \cdot \gamma e) + f(s), \quad u \in \Phi(s), \\ \text{or } n\partial_t \Phi^{-1}(u) &= \mu^{-1} \operatorname{div}(\nabla u + k(\Phi^{-1}(u)) \cdot \gamma e) + f(\Phi^{-1}(u)). \end{aligned} \quad (7)$$

Conversely, the inverse Kirchhoff transformation  $\kappa^{-1} : u \mapsto p$  can be defined as follows (cf. [13])

$$p = \kappa^{-1}(u) := \begin{cases} -h(u) & \text{if } u \leq 0, \\ \frac{u_+}{k(1)} & \text{else} \end{cases}, \quad h(u) := \int_{u_-}^0 \frac{1}{k(\Phi^{-1}(\tau))} d\tau. \quad (8)$$

With the global pressure formulation (7) we are now prepared to state our fully coupled system.

## 2.1 Fully coupled system for unsaturated flow in porous media with infiltration and Signorini-type outflow conditions

Let us define  $\Omega_T = \Omega \times [0, T]$ ,  $\Omega_0 = \Omega \times \{0\}$  and  $\Sigma_{i,t} = \Sigma_i \times [0, t]$  for  $t \in [0, T]$  with  $i = \text{in, out, } N$ . With given initial values  $s_0$  and  $w_0$ , we get the following system of differential equations for the unknown global pressure  $u$  and surface water height  $w$

$$\begin{aligned}
u &\in \Phi(s), \quad \text{and } \partial_t s = \operatorname{div}(\nabla u + k(s) \cdot \mathbf{e}) + f(s) && \text{in } \Omega_T, \\
-(\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &\geq 0 && \text{on } \bar{\Sigma}_{\text{out},T} \cup \Sigma_{N,T}, \\
u &\leq \tilde{\Phi}(1) \quad \text{and } k^2(s)s - k^2(1)1 \leq 0 && \text{on } \Sigma_{\text{out},T}, \\
0 &\leq -((\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu}) \cdot (k^2(s)s - k^2(1)1) && \text{on } \Sigma_{\text{out},T}, \\
(\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 && \text{on } \Sigma_{N,T}, \\
(\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= g(u, w) && \text{on } \Sigma_{\text{in},T}, \\
s &= s_0 && \text{in } \Omega_0, \\
\partial_t w &= (r - g(u, w)) && \text{in } \Sigma_{\text{in},T}, \\
w &= w_0 && \text{on } \Sigma_{\text{in},0},
\end{aligned} \tag{9}$$

where

$$g(u, w) = -\frac{u_+}{k(1)c} + \frac{w}{c} + \frac{h(u)\psi(w)}{c}, \quad \text{and } \psi(w) = \min \left\{ 1, \left( \frac{w}{\sigma} \right)_+ \right\}.$$

As the mappings  $s \mapsto k^2(s)s$  and  $s \mapsto p_c(s)$  are both monotone functions, and as  $p_{\text{gas}} = 0$  and  $\tilde{p}_c(1) = 0$ , the outflow conditions stated in (2) and the outflow conditions prescribed in (9) are formally equivalent.

It is also possible to allow  $\tilde{p}_c(1) < 0$  as it is the case, e.g., for Brooks–Corey parameter functions [16]. But then the outflow conditions in (2) and the outflow conditions in (9) are no longer equivalent. We would have to deal either with the conditions in (2) or with similar conditions formed by replacing  $p$  by  $u$  in (2) as done, e.g., in [10]. In such case it would be very difficult to give a meaning to the traces in the last equality in (2), as it demands an a priori estimate for  $\|\operatorname{div}(\mathbf{q})\|_{L^2(\Omega)}$ , which in turn necessitates an estimate for  $\|\partial_t s\|_{L^2(\Omega)}$ . In order to solve this problem, one can either assume  $\|\partial_t s\|_{L^2(\Omega)} \leq C$  or  $\Phi'(s) \geq c_1 > 0$ . If one makes one of these assumptions, it is also possible to show an existence result for the system with original outflow conditions (2), based on the results in [34]. Otherwise, it is only feasible to derive a weighted  $L^2$ -estimate for  $\operatorname{div}(\mathbf{q})$  [37] and the original outflow conditions have to be replaced by some surrogate ones as suggested in [37] and (9).

## 2.2 Assumptions on the data

In order to guarantee among others the well-definedness of the boundary conditions, the coefficient functions have to fulfill several assumptions. Additional assumptions are needed for the proof of our existence result (see Theorem 3.1).

The relative permeability and capillary pressure functions are supposed to fulfill the following assumption.

soil type	$k(1)\gamma/\mu$ [cm/day]	$\alpha$ [cm $^{-1}$ ]	$l$ [-]
Hygiene sandstone	108.0	0.0079	10.4
Touchet Silt Loam G.E.3	303.0	0.005	7.09
Silt Loam G.E.3	4.96	0.00423	2.06
Guelph Loam (drying)	31.6	0.0115	2.03
Guelph Loam (wetting)	-	0.02	2.76
Beit Netofa Clay	0.082	0.00152	1.17

Table 1: Typical values for  $k(1)\gamma/\mu$ ,  $\alpha$  and  $l$  used for the van Genuchten parametrization [40] for five different soil types.

**Assumption 2.1.** *The permeability*

$$k \in C^1([0, 1], [0, \infty)) \text{ is monotonically non-decreasing}$$

and  $k(s) = 0$  for  $s = 0$ . The capillary pressure  $p_c : (0, 1] \rightarrow \{0, 1\}^{\mathbb{R}}$  is a monotone graph given by a function

$$\tilde{p}_c \in C^1((0, 1), \mathbb{R}), \text{ monotonically increasing.}$$

We set  $p_c(s) = \{\tilde{p}_c(s)\}$  for  $s \in (0, 1)$  and  $p_c(1) = [\tilde{p}_c(1), \infty)$ . Furthermore, we suppose  $\tilde{p}_c(1) = 0$  (hydrophilic case). Finally, we assume that with a constant  $c_0 > 0$  the following conditions hold:

$$\begin{aligned} \partial_s p_c(s) &\geq 1/c_0 & \forall s \in (0, 1), \\ |\partial_s k(s)|^2 &\leq c_0 k(s) & \forall s \in (0, 1), \\ k(s)|p_c(s)| + \sqrt{k(s)}\partial_s p_c(s) &\leq c_0 & \forall s \in (0, 1/2), \\ (1-s)\sqrt{\partial_s p_c(s)} &\leq c_0 & \forall s \in (1/2, 1). \end{aligned} \tag{10}$$

In order to guarantee that the condition on the infiltration boundary is well-defined, we additionally presume that the coefficient functions are chosen in such a way that

$$u_- \in L^\infty(\Omega_T) \implies p_- \in L^\infty(\Omega_T). \tag{11}$$

Taking Assumption 2.1 for granted, we can prove  $u_- \in L^\infty(\Omega_T)$ . Together with (11) this yields the well-definedness of the conditions on the infiltration boundary.

To justify the above assumptions from a physical viewpoint, we check if they can be fulfilled for the van Genuchten model [40] of the coefficient functions

$$k(s) = s^{1/2} \cdot (1 - (1 - s^{1/m})^m)^2, \quad p_c(s) = -\frac{1}{\alpha}(s^{-1/m} - 1)^{1/l},$$

with  $m = 1 - 1/l$ . In Table 1 some typical values for  $k(1)\gamma/\mu$  — the hydraulic conductivity at full saturation — and two variable parameters  $\alpha$  and  $l$  for five different soil types are listed. The values are extracted from [40]. Starting with the implication (11), we see that this holds true for  $l < 3$ . Concerning estimates (10), we observe that they are all fulfilled except for the second one. This is due to the fact that the description of the relative permeability  $k(s)$  proposed by

van Genuchten is bounded for  $s = 1$ , but on the other hand its derivative blows up. However, since unbounded derivatives of  $k$  are hydrologically unrealistic, we can modify  $k(s)$  for  $s$  near 1 so that all estimates (10) are satisfied. We emphasize the importance of the third inequality in (10), which guarantees that  $k$  declines fast enough for small  $s$  to absorb the pressure  $p_c$ , which goes to  $-\infty$  in this case.

Another parametrization of  $k(s)$  and  $p_c(s)$  is given by Brooks and Corey [16], see also [17] and [40]. One can check that all estimates (10) are satisfied for this parametrization, however, with regard to Assumption 2.1, we always have  $\tilde{p}_c(1) < 0$ . Moreover, it is a characteristic property of the Brooks–Corey functions that the crucial implication (11) is never fulfilled. On the contrary, the inverse Kirchhoff transformation (8) is always ill-posed around the minimal global pressure  $u_c = \Phi(0)$  with  $\kappa^{-1}(u_c) = -\infty$ , cf. [10, Sec. 1.3]. With regard to the analysis in physical variables, we can consider regularizations of Brooks–Corey functions which exhibit the non-degenerate case  $k(s) > c_0 > 0$ , for which implication (11) is satisfied, cf. [10, Sec. 1.4.3] and Section 4. However, even though the numerics for the Richards equation is challenging for the degenerate Brooks–Corey functions, we use them in our numerical examples in Section 5 analogue to [13].

Concerning the source term, we assume that  $f : \Omega_T \times [0, 1] \rightarrow \mathbb{R}$  is bounded, continuously differentiable in all entries, and monotonically decreasing in  $s$ . Moreover,  $f$  shall satisfy  $f(x, t, 1) \leq 0$  and  $f(x, t, 0) \geq 0$  for all  $(x, t) \in \Omega_T$ . Eventually, we suppose that  $s \mapsto f(x, t, s)$  is affine on  $(0, \lambda)$  for small  $\lambda$ . Furthermore, we assume that  $r : \Sigma_{\text{in}, T} \rightarrow \mathbb{R}^+$  is bounded and Lipschitz continuous. The initial conditions shall be given by functions  $s_0 : \Omega \rightarrow [0, 1]$  and  $w_0 : \Sigma_{\text{in}} \rightarrow \mathbb{R}^+$ , with  $s_0 \in L^\infty(\Omega)$  and  $w_0 \in L^\infty(\Sigma_{\text{in}})$  and

$$s(x, 0) = s_0(x) \quad \text{a.e. in } \Omega \quad \text{and} \quad w(x, 0) = w_0(x) \quad \text{a.e. on } \Sigma_{\text{in}}.$$

We impose the following compatibility condition for  $s_0$ :

$$\begin{aligned} \exists p_0 : \Omega \rightarrow \mathbb{R}, \quad p_0 \in p_c(s_0) \quad \text{a.e. on } \{k(s_0) > 0\}, \\ p_0 \in H^{1,2}(\Omega) \cap L^\infty(\Omega), \quad s_0|_{\Sigma_{\text{in}}} > 0. \end{aligned}$$

### 3 Existence result for the degenerate coupled system

The goal of this section is to prove an existence result for system (9). Unfortunately, (9) exhibits some unpleasant properties. First, the system is degenerate in the sense that  $k(s) \rightarrow 0$  for  $s \rightarrow 0$  and that the derivative  $p'_c(s)$  may be unbounded in the neighbourhood of  $s = 0$  and  $s = 1$ . As a result, the solution of the Richards equation lacks regularity, making the treatment of the boundary conditions challenging, as it is difficult to give sense to the appearing traces of the functions. Moreover, we deal with a coupled system due to the boundary conditions at the infiltration boundary. To achieve our existence result, we thus first regularize the coefficient functions  $k(s)$  and  $p_c(s)$  with a small regularization parameter  $\delta$  in order to obtain a non-degenerate system. Afterwards, we fix  $\delta$  and decouple the system. Existence of solutions for the decoupled sub-problems

then follows by standard arguments. By applying Banach's fixed point theorem, we get a unique solution  $(u_\delta, w_\delta)$  for the regularized coupled system system, for fixed  $\delta$ . The crucial part when proving a priori estimates independent of  $\delta$  is the derivation of a maximum principle for  $u_\delta$  and  $w_\delta$ . This can be achieved by considering classical solutions of a regularized parabolic system. Finally, we pass to the limit in  $\delta$  and get a solution for our original problem (9). Our main result summarizes as follows.

**Theorem 3.1** (Existence of a solution pair  $(u, w)$  of system (9)). *Let the assumptions from Section 2 be fulfilled and let the pair  $(u_\delta, w_\delta)$  be a sequence of regularized solutions in  $L^\infty(\Omega_T) \cap L^2((0, T), H^{1,2}(\Omega)) \times C^0([0, T], L^1(\Sigma_{in})) \cap L^\infty((0, T), L^2(\Sigma_{in}))$ . Then there exists a subsequence  $(u_\delta, w_\delta)$  and a solution pair  $(u, w) \in L^2((0, T), H^{1,2}(\Omega)) \times C^0([0, T], L^1(\Sigma_{in}))$  such that*

$$\begin{aligned} u_\delta &\rightharpoonup u \text{ in } L^2((0, T), H^{1,2}(\Omega)), \\ w_\delta &\rightarrow w \text{ in } C^0([0, T], L^1(\Sigma_{in})), \end{aligned}$$

and  $(u, w)$  are solutions of (9) in the distributional sense. Additionally,  $w \geq 0$  holds a.e. on  $\Sigma_{in}$  for all  $t \in [0, T]$ .

In the following subsections we will prove Theorem 3.1 as indicated above.

### 3.1 Regularization of the coupled problem

At first, we regularize the coefficient functions.

**Assumption 3.2** (Regularized coefficient functions). *The regularized coefficient functions should fulfill the following conditions:*

$$k_\delta \in C^1([0, 1], (0, \infty)) \quad \text{and} \quad \rho_\delta \in C^0([0, 1], \mathbb{R}) \text{ piecewise } C^1$$

are monotonically increasing. For  $\delta \rightarrow 0$  we have  $k_\delta \searrow k$  uniformly on  $[0, 1]$ ,  $\rho_\delta \rightarrow p_c$  uniformly on compact subsets of  $(0, 1)$  and  $k_\delta(0) = \delta^2$ . We suppose that  $\bigcup_\delta \rho_\delta([0, 1]) = \mathbb{R}$  and, for simplicity, that  $p_c(1/2) = \rho_\delta(1/2)$  holds. Finally,  $k_\delta$  and  $\rho_\delta$  shall fulfill the inequalities of Assumption 2.1.

The Assumptions 2.1 and 3.2 allow the definition of a regularized global pressure via the Kirchhoff transformation

$$u_\delta(x, t) = \Phi_\delta(s) = \int_0^{\rho_\delta(s)} k_\delta(\rho_\delta^{-1}(q)) dq,$$

and guarantee that  $\Phi_\delta$  is bounded from below and that  $\Phi_\delta \rightarrow \Phi$  uniformly on compact subsets of  $[0, 1]$ . We set  $u_c = \Phi(0)$  as the minimal global pressure.

**Example 3.3.** *Let Assumption 2.1 be fulfilled and suppose that  $c_1 s^2 \leq k(s) \leq c_2 s^2$  on  $(0, 1)$  for constants  $0 < c_1 \leq c_2$ . Then the following regularization fulfills Assumption 3.2.*

$$k_\delta(s) := \delta^2 + k(s) \quad \forall s \in [0, 1], \quad \rho_\delta(s) := \begin{cases} p_c(\delta) + \frac{s-\delta}{\delta} & \forall s \in [0, \delta], \\ p_c(s) & \forall s \in (\delta, 1-\delta], \\ p_c(1-\delta) + \frac{s-(1-\delta)}{\delta^2} & \forall s \in (1-\delta, 1]. \end{cases}$$

Note that all inequalities (10), which have to be satisfied by the regularized coefficient functions, except for inequality (10)<sub>3</sub>, are needed for the a priori estimates. If (10)<sub>3</sub> is not fulfilled, we would still be able to prove convergence of the approximate solution pair to a limit  $(u, w)$ , but this limit would not necessarily solve our problem.

Using the regularized coefficient functions, we get the regularized Richards equation

$$\partial_t s_\delta = \operatorname{div}(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) + f_\delta, \quad u_\delta = \Phi_\delta(s_\delta),$$

with the pressure  $p_\delta = \rho_\delta(s_\delta)$  and  $f_\delta = f(x, t, s_\delta(x, t))$ . In a straightforward manner we obtain the conditions on the Neumann and the infiltration boundary:

$$\begin{aligned} (\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= g_\delta(u_\delta, w_\delta) && \text{on } \Sigma_{\text{in}, T} \\ (\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 && \text{on } \Sigma_{N, T}. \end{aligned}$$

Inspired by [37], we impose the following nonlinear Robin condition on the outflow boundary:

$$-(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} = \frac{1}{\delta} k_\delta(s_\delta)(p_\delta)_+. \quad (12)$$

In this way, we keep the idea behind the original outflow conditions, as water can flow out only if  $p \geq 0$  and if  $p > 0$  we have a very large outflow due to the scaling  $1/\delta$ . On the other hand, boundary condition (12) is more handy than the original one, as we do not have to deal with a variational inequality. Rewriting (12) with  $\mathbf{q}_\delta = -(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e})$  one gets

$$\mathbf{q}_\delta \cdot \boldsymbol{\nu} = k_\delta(s_\delta) \frac{(p_\delta)_+ - 0}{\delta},$$

which, similar to our condition on the infiltration boundary, can be interpreted as a pressure difference between the pressure inside the porous medium and the pressure outside  $p_{\text{gas}} = 0$ . The initial conditions are replaced by

$$s_\delta(x, 0) = s_0^\delta(x) := \begin{cases} \rho_\delta^{-1}(p_0(x)) & \text{if } s_0 > 1/2, \\ s_0(x) & \text{if } s_0 \leq 1/2, \end{cases}$$

for all  $x \in \Omega$ , guaranteeing  $\Phi_\delta(s_0^\delta) \in H^{1,2}(\Omega)$ .

We get the following regularized system for the unknowns  $u_\delta$  and  $w_\delta$

$$\begin{aligned} u_\delta &= \Phi_\delta(s_\delta), \quad \partial_t s_\delta = \operatorname{div}(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) + f(s_\delta) && \text{in } \Omega_T, \\ -(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= \frac{1}{\delta} k_\delta(s_\delta)(p_\delta)_+, && \text{on } \Sigma_{\text{out}, T}, \\ (\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 && \text{on } \Sigma_{N, T}, \\ (\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= g_\delta(u_\delta, w_\delta) && \text{on } \Sigma_{\text{in}, T}, \\ s_\delta(x, 0) &= s_0^\delta := \begin{cases} \rho_\delta^{-1}(p_0(x)) & \text{if } s_0(x) > \frac{1}{2}, \\ s_0(x) & \text{if } s_0(x) \leq \frac{1}{2}, \end{cases} && \text{in } \Omega_0, \\ \partial_t w_\delta &= (r - g_\delta(u_\delta, w_\delta)) && \text{in } \Sigma_{\text{in}, T}, \\ w_\delta &= w_0 && \text{on } \Sigma_{\text{in}, 0}, \end{aligned} \quad (13)$$

where

$$g_\delta(u_\delta, w_\delta) = -\frac{(u_\delta)_+}{k_\delta(1)c} + \frac{w_\delta}{c} + \frac{h_\delta(u_\delta)\psi(w_\delta)}{c},$$

$$h_\delta(u_\delta) = \int_{(u_\delta)_-}^0 \frac{1}{k_\delta(\Phi_\delta^{-1}(\tau))} d\tau, \quad \psi(w_\delta) = \min \left\{ 1, \left( \frac{w_\delta}{\sigma} \right)_+ \right\}.$$

### 3.2 Existence of a unique solution for fixed $\delta$

Let  $\delta$  be fixed. To prove the existence of a unique solution pair  $(u_\delta, w_\delta) \in L^2((0, T), H^{1,2}(\Omega)) \times C^0([0, T], L^1(\Sigma_{\text{in}}))$  of (9), we take the complete metric space  $M = C^0([0, T], L^1(\Sigma_{\text{in}}))$  with the metric  $\|w_1 - w_2\|_M = \max_{0 \leq t \leq T} \int_{\Sigma_{\text{in}}} |w_1(t) - w_2(t)| dx$ , construct a mapping  $A : M \rightarrow M$  and show that this mapping is  $k$ -contractive. The fixed point of the mapping  $A$  is the solution of the ODE-subproblem. Inserting this solution as data into the subproblem of the Richards equation results in the solution pair of the whole system. For the sake of readability, we omit the index  $\delta$  in this subsection. At first, we define the two subproblems.

**Definition 3.4** (Subproblem of the Richards equation). *Let  $w_{j-1} \in M \cap L^\infty((0, T), L^2(\Sigma_{\text{in}}))$  be given for  $j = 1, 2, \dots$ . We characterize  $u_j$  as the solution of the problem*

$$\begin{aligned} \partial_t \Phi^{-1}(u_j) - \operatorname{div}(\nabla u_j + k(\Phi^{-1}(u_j)) \cdot \mathbf{e}) &= f(\Phi^{-1}(u_j)) && \text{in } \Omega_T, \\ -(\nabla u_j + k(\Phi^{-1}(u_j)) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= \frac{1}{\delta} k(\Phi^{-1}(u_j)) (p_j)_+ && \text{on } \Sigma_{\text{out}, T}, \\ (\nabla u_j + k(\Phi^{-1}(u_j)) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 && \text{on } \Sigma_{N, T}, \\ (\nabla u_j + k(\Phi^{-1}(u_j)) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= g(u_j, w_{j-1}) && \text{on } \Sigma_{\text{in}, T}, \\ \Phi^{-1}(u_j)(x, 0) &= s_0 && \text{in } \Omega_0, \end{aligned} \tag{14}$$

with

$$g(u_j, w_{j-1}) = -\frac{(u_j)_+}{k(1)c} + \frac{w_{j-1}}{c} + \frac{h(u_j)\psi(w_{j-1})}{c},$$

$$h(u_j) = \int_{(u_j)_-}^0 \frac{1}{k(\Phi^{-1}(\tau))} d\tau, \quad \psi(w_{j-1}) = \min \left\{ 1, \left( \frac{w_{j-1}}{\sigma} \right)_+ \right\}.$$

**Definition 3.5** (Weak formulation of the subproblem of the Richards equation). *We say that  $u_j \in L^2((0, T), H^{1,2}(\Omega))$  is a weak solution of problem (14), if  $\Phi^{-1}(u_j) \in L^2((0, T), L^2(\Omega))$  and*

$$\begin{aligned} 0 &= \int_{\Omega} \Phi^{-1}(u_j)(0, x) \varphi(0, x) dx + \int_{\Omega_T} \Phi^{-1}(u_j) \partial_t \varphi \\ &\quad - \int_{\Omega_T} (\nabla u_j + k(\Phi^{-1}(u_j)) \cdot \mathbf{e}) \cdot \nabla \varphi + \int_{\Omega_T} f(\Phi^{-1}(u_j)) \varphi \\ &\quad - \int_{\Sigma_{\text{out}, T}} \frac{1}{\delta} (k(\Phi^{-1}(u_j))(p_j)_+) \varphi + \int_{\Sigma_{\text{in}, T}} g(u_j, w_{j-1}) \varphi \end{aligned} \tag{15}$$

for all test functions  $\varphi \in L^2((0, T), H^{1,2}(\Omega))$  with  $\partial_t \varphi \in L^2((0, T), L^2(\Omega))$  and  $\varphi(T) = 0$ .

**Lemma 3.6** (Existence of a unique weak solution). *Let  $w_{j-1} \in M \cap L^\infty((0, T), L^2(\Sigma_{in}))$  be given for  $j \in \mathbb{N}$ , and let the assumptions from Section 2 and Subsection 3.1 be fulfilled. Then there exists a unique weak solution  $u_j \in L^2((0, T), H^{1,2}(\Omega))$  in the sense of Definition 3.5.*

*Proof.* A proof can be found, e.g., in [30].  $\square$

**Definition 3.7** (ODE subproblem). *Let  $w_{j-1} \in M \cap L^\infty((0, T), L^2(\Sigma_{in}))$  and  $u_j \in L^\infty(\Omega_T) \cap L^2((0, T), H^{1,2}(\Omega))$  be given for  $j = 1, 2, \dots$ . We characterize  $w_j$  as the solution of the problem*

$$\begin{aligned} \partial_t w_j &= r - g(u_j, w_j, w_{j-1}) && \text{in } \Sigma_{\text{in}, T} \\ w_j &= w_0 && \text{on } \Sigma_{\text{in}, 0}, \end{aligned} \tag{16}$$

with

$$\begin{aligned} g(u_j, w_j, w_{j-1}) &= -\frac{(u_j)_+}{k(1)c} + \frac{w_j}{c} + \frac{h(u_j)\psi(w_{j-1})}{c}, \\ h(u_j) &= \int_{(u_j)_-}^0 \frac{1}{k(\Phi^{-1}(\tau))} d\tau, \quad \psi(w_{j-1}) = \min \left\{ 1, \left( \frac{w_{j-1}}{\sigma} \right)_+ \right\}. \end{aligned}$$

**Lemma 3.8** (Existence of a unique solution of the ODE subproblem). *Let  $w_{j-1} \in M \cap L^\infty((0, T), L^2(\Sigma_{in}))$  and  $u_j \in L^\infty(\Omega_T) \cap L^2((0, T), H^{1,2}(\Omega))$  be given for  $j \in \mathbb{N}$ , and let the assumptions from Section 2 and Subsection 3.1 be fulfilled. Then there exists a unique solution  $w_j \in M$  of the initial value problem (16).*

*Proof.* It can be easily proved that the right-hand side of the ODE (16) fulfills the assumptions of the Carathéodory's existence theorem for ODEs, which in turn yields the claim.  $\square$

We proceed with some estimates for the functions  $u_j$  and  $w_j$ .

**Corollary 3.9** (A priori bounds for fixed  $\delta$ ). *Under the assumptions from Section 2 and Subsection 3.1 and with a constant  $C > 0$ , the following estimates for fixed  $\delta > 0$  and arbitrary index  $j$  hold true:*

1.  $s_\delta \in [0, 1]$ ,  $s_\delta \in L^2([0, T], H^{1,2}(\Omega))$ ,  $\int_{\Omega_T} |\partial_t s_\delta|^2 \leq C$ ,
2.  $u_\delta \in L^\infty(\Omega_T) \cap L^2([0, T], H^{1,2}(\Omega))$ ,  $\int_{\Omega_T} |\partial_t u_\delta|^2 \leq C$ ,
3.  $\|\operatorname{div}(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e})\|_{L^2(\Omega_T)} \leq C$ ,
4.  $p_\delta(s_\delta) \in L^\infty(\Omega_T)$ ,  $p'_\delta(s_\delta) \in L^\infty(\Omega_T)$ ,
5.  $h_\delta(u_\delta) \in L^\infty(\Sigma_T)$ ,
6.  $\Phi_\delta(s_\delta)$  is Lipschitz continuous in  $s_\delta$ .

*Proof.* This is a simple conclusion from the estimates for arbitrary  $\delta$ , which will be derived in Subsection 3.3.  $\square$

Now we can define our mapping  $A : M \rightarrow M$  as a composition of two mappings  $B : M \rightarrow L^2((0, T), H^{1,2}(\Omega))$  and  $D : M \times L^2((0, T), H^{1,2}(\Omega)) \rightarrow M$  with  $u_j := Bw_{j-1}$  and  $w_j := D(w_{j-1}, u_j)$ , where  $B$  and  $D$  are defined by the subproblems (14) and (16). So  $A$  can be defined as  $w_j = D(w_{j-1}, Bw_{j-1}) =: Aw_{j-1}$ . We now prove the  $k$ -contractivity of  $A$ .

**Proposition 3.10** (Estimate for the boundary conditions and the saturation). *Let  $w_{j-1}, \tilde{w}_{j-1} \in M \cap L^\infty((0, T), L^2(\Sigma_{in}))$  be two solutions of the initial value problem (16) and  $u_j, \tilde{u}_j \in L^\infty(\Omega_T) \cap L^2((0, T), H^{1,2}(\Omega))$  the associated solutions of (14) for  $j = 1, 2, \dots$ . Furthermore, let the assumptions from Section 2 and Subsection 3.1 be fulfilled. Then for  $0 \leq t \leq T$ ,  $s_j = \Phi^{-1}(u_j)$ ,  $\tilde{s}_j = \Phi^{-1}(\tilde{u}_j)$  and a constant  $C > 0$  the following estimate holds true:*

$$\begin{aligned} & \int_{\Omega} |s_j(\cdot, t) - \tilde{s}_j(\cdot, t)| + \int_{\Sigma_{in,t}} \frac{1}{k(1)c} |(u_j)_+ - (\tilde{u}_j)_+| + \frac{|\psi(\tilde{w}_{j-1})|}{c} |h(u_j) - h(\tilde{u}_j)| \\ & + \int_{\Sigma_{out,t}} \frac{1}{\delta} |k(s_j)(p_+)(s_j) - k(\tilde{s}_j)(p_+)(\tilde{s}_j)| \leq C \|w_{j-1} - \tilde{w}_{j-1}\|_{L^1(\Sigma_{in,t})}. \end{aligned} \quad (17)$$

The constant  $C$  can be specified as  $C = c^{-1}(1 + \|h(u_j)\|_{L^\infty(\Sigma_{in,t})} L_\psi)$  where  $L_\psi$  is the Lipschitz constant of  $\psi$ .

*Proof.* To prove this estimate, we use duality techniques, which is done for a similar case in detail, e.g., in [23] and for the case here in [38]. In short, we put all derivatives onto the test function  $\varphi$  and solve classically a dual problem in  $\varphi$ , which, loosely speaking, subtracts the boundary conditions appearing in the weak formulation, and adds them in absolute values, arranged with the right algebraic sign for the estimate.  $\square$

From the integral formulation of the ODE we get with straightforward modifications

$$\begin{aligned} \int_{\Sigma_{in}} |w_j - \tilde{w}_j| & \leq \frac{1}{k(1)c} \int_{\Sigma_{in,t}} |(u_j)_+ - (\tilde{u}_j)_+| + \frac{1}{c} \int_{\Sigma_{in,t}} |w_j - \tilde{w}_j| \\ & + C \int_{\Sigma_{in,t}} |w_{j-1} - \tilde{w}_{j-1}| + \frac{1}{c} \int_{\Sigma_{in,t}} \psi(\tilde{w}_{j-1}) |h(u_j) - h(\tilde{u}_j)|, \end{aligned} \quad (18)$$

with  $C$  having been defined in Proposition 3.10. This estimate (18) together with (17) give

$$\int_{\Sigma_{in}} |Aw_{j-1} - A\tilde{w}_{j-1}| = \int_{\Sigma_{in}} |w_j - \tilde{w}_j| \leq \frac{1}{c} \int_{\Sigma_{in,t}} |w_j - \tilde{w}_j| + \hat{C} \int_{\Sigma_{in,T}} |w_{j-1} - \tilde{w}_{j-1}|,$$

where  $\hat{C} = 2C$ . The application of Gronwall's lemma yields

$$\begin{aligned} \int_{\Sigma_{in}} |w_j - \tilde{w}_j| & \leq \hat{C} e^{\frac{t}{c}} \int_{\Sigma_{in,T}} |w_{j-1} - \tilde{w}_{j-1}| \leq \hat{C} e^{\frac{T}{c}} \int_{\Sigma_{in,T}} |w_{j-1} - \tilde{w}_{j-1}| \\ & \leq T \hat{C} e^{\frac{T}{c}} \max_{0 \leq t \leq T} \int_{\Sigma_{in}} |w_{j-1} - \tilde{w}_{j-1}|, \end{aligned}$$

which results in

$$\|Aw_{j-1} - A\tilde{w}_{j-1}\|_M \leq T \hat{C} e^{\frac{T}{c}} \|w_{j-1} - \tilde{w}_{j-1}\|_M. \quad (19)$$

For  $T \cdot \hat{C} \cdot e^{\frac{T}{c}} = \theta < 1$ ,  $A$  is obviously  $k$ -contractive. Since  $w_0 = \tilde{w}_0$ , (19) provides the uniqueness of the solutions  $w_j$  and  $u_j$ , too. With Banach Fixed Point Theorem we get existence of a unique solution pair  $(u_\delta, w_\delta)$  of the regularized system (13). We summarize this result in the following theorem.

**Theorem 3.11.** *For each  $\delta > 0$  there exists a unique solution pair  $u_\delta \in L^\infty(\Omega_T) \cap L^2((0, T), H^{1,2}(\Omega))$  and  $w_\delta \in M \cap L^\infty((0, T), L^2(\Sigma_{in}))$  of the regularized system (13). Furthermore, we have  $w_\delta \geq 0$  a.e. on  $\Sigma_{in}$  for all  $t \in [0, T]$  and each  $\delta > 0$ .*

*Proof.* It remains to prove that  $w_\delta \geq 0$  a.e. on  $\Sigma_{in}$  for all  $t \in [0, T]$ . We discretize the regularized system (13) explicitly in time, where we omit the index  $\delta$  again. By  $w^n$  and  $u^n$  we denote the time-discrete approximation of  $w(\cdot, t_n)$  and  $u(\cdot, t_n)$  respectively for a time step  $t_n \in [0, T]$ ,  $n = 0, 1, \dots, N$ ,  $t_0 := 0$ ,  $t_N := T$  and a time step size  $\tau = t_{n+1} - t_n$ . We show by means of mathematical induction that  $w^n(x) \geq 0$  a.e. on  $\Sigma_{in}$  for all  $n \in \mathbb{N}$ . As  $w^0(x) \geq 0$  by Assumptions on the data 2.2, we prove now the induction step  $w^n(x) \geq 0 \Rightarrow w^{n+1}(x) \geq 0$  a.e. on  $\Sigma_{in}$ . To this end, we consider for fixed  $x \in \Sigma_{in}$  the scheme

$$\frac{w^{n+1}(x) - w^n(x)}{\tau} = r(x, t_n) + \frac{u^n(x)_+}{k(1)c} - \frac{w^n(x)}{c} - \frac{h(u^n(x)) \cdot \min\{1, (\frac{w^n(x)}{\sigma})_+\}}{c} \quad (20)$$

for the solution of the ODE (16) on  $\Sigma_{in,T}$ . Please note that  $u_\delta(x)$  is a solution of a non-degenerate elliptic problem and therefore regular enough to be evaluated on  $\Sigma_{in}$ . Let  $w^n(x) \geq 0$  a.e. on  $\Sigma_{in}$  be fulfilled. Rearranging (20) leads to

$$w^{n+1}(x) = \left(1 - \frac{\tau}{c}\right) w^n(x) + \underbrace{\tau \left(r(x, t_n) + \frac{u^n(x)_+}{k(1)c}\right)}_{\geq 0} - \tau \left(\frac{h(u^n(x))\psi(w^n(x))}{c}\right).$$

Next we check which values  $\tau$  guarantee that  $w^{n+1}(x) \geq 0$ , too. By simple modifications one gets that the estimate

$$\tau \leq \min \left\{ c, \frac{\sigma}{\frac{\sigma}{c} - r(x, t_n) + \frac{h(u^n(x))}{c}}, \frac{c}{1 + \frac{h(u^n(x))}{\sigma}} \right\} \quad \text{for } x \in \Sigma_{in} \text{ a.e.} \quad (21)$$

has to hold true if the second and third entry in the braces are nonnegative—otherwise these terms do not appear in (21). Now, if we let  $\tau$  go to zero and use the global estimate

$$\tau \leq \underset{t \in [0, T]}{\text{ess min}} \left\{ c, \frac{\sigma}{\frac{\sigma}{c} - r(x, t) + \frac{h(u(x, t))}{c}}, \frac{c}{1 + \frac{h(u(x, t))}{\sigma}} \right\} \quad \text{for } x \in \Sigma_{in} \text{ a.e.,}$$

we obtain  $w_\delta \geq 0$  for all  $\delta > 0$ . Note that  $r \in C^0(\bar{\Sigma}_{in,T})$  (2.2) and  $u \in L^\infty(\Sigma_{in,T})$  imply that the introduced explicit Euler method converges pointwise a.e. on  $\Sigma_{in}$ .  $\square$

We remark that estimate (21) can in principle also be used for numerical simulations in order to guarantee that the numerical solution remains positive. For each time  $t_n$ , using an approximation of  $u(\cdot, t_n)$  in (21), an upper bound  $Dt$  for  $\tau$  can be computed with (21) ensuring that the approximation of  $w(\cdot, t_{n+1})$  remains nonnegative for all  $\tau \leq Dt$ . However, numerical tests show that this bound is quite strict, see Section 5.

### 3.3 A priori bounds

In this subsection we derive a priori bounds for the solution pair  $(u_\delta, w_\delta)$  of the regularized system (13) that are independent of  $\delta$ . Let the assumptions of Section 2 and Subsection 3.1 be fulfilled.

**Lemma 3.12.** *There exists  $C < \infty$  independent of  $\delta > 0$  such that, for all  $\delta > 0$ ,*

$$\int_{\Omega_T} k_\delta \rho'_\delta |\partial_t s_\delta|^2 \leq C \quad \text{and} \quad \int_{\Omega_T} k_\delta(s_\delta) |\operatorname{div}(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e})|^2 \leq C.$$

*Proof.* The first estimate is obtained by testing the weak formulation (15) with  $\partial_t u$ , using Assumption 3.2 and writing the integrands of the integrals on the boundary as total time derivatives. The second estimate follows directly from the first one. For a detailed exposition of the proof we refer to [38] and, in the case of Dirichlet boundary conditions on the infiltration boundary, to [37].  $\square$

**Theorem 3.13** (Maximum principle). *There exists  $C < \infty$  independent of  $\delta$  such that, for all  $\delta > 0$*

$$\|(u_\delta)_+\|_{L^\infty(\Sigma_{in,T})} + \|w_\delta\|_{L^\infty(\Sigma_{in,T})} \leq C, \quad \text{and} \quad \|u_\delta\|_{L^\infty(\Omega_T)} \leq C.$$

Furthermore, there exists  $p_{max} < \infty$  independent of  $\delta$  such that  $p_\delta(x, t) \leq p_{max}$  for almost all  $(x, t) \in \Omega_T$  and all  $\delta > 0$ .

For the derivation of a maximum principle for the solution pair  $(u_\delta, w_\delta)$ , we seize ideas of [24]. We can obtain  $L^\infty$ -estimates independent of  $\delta$  in the case where the pair  $(u_\delta, w_\delta)$  can be approximated by smooth functions  $(u_\delta^\varepsilon, w_\delta^\varepsilon)$  which are solutions of a regularized parabolic system. For the sake of simplicity we assume that the boundary of  $\Omega$  is of class  $C^2$ . If  $\partial\Omega$  is not smooth enough one can proceed as in [24] and require that the elevation of the surface can be smoothly and periodically extended to the whole  $\mathbb{R}^2$ . Furthermore, we presume for the proof of Theorem 3.13 that  $k_\delta \in C^2([0, 1], (0, \infty))$  and  $\rho_\delta \in C^2([0, 1], \mathbb{R})$ , otherwise an additional regularization step has to be performed. We consider the parabolic system

$$\begin{aligned} u_\delta^\varepsilon &= \Phi_\delta(s_\delta^\varepsilon), \quad \partial_t s_\delta^\varepsilon = \operatorname{div}(\nabla u_\delta^\varepsilon + k_\delta(s_\delta^\varepsilon) \cdot \mathbf{e}) + f^\varepsilon(s_\delta^\varepsilon) && \text{in } \Omega_T, \\ -(\nabla u_\delta^\varepsilon + k_\delta(s_\delta^\varepsilon) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= \frac{1}{\delta} k_\delta(s_\delta^\varepsilon) \xi^\varepsilon(p_\delta^\varepsilon) \chi_\varepsilon, && \text{on } \Sigma_{out,T}, \\ (\nabla u_\delta^\varepsilon + k_\delta(s_\delta^\varepsilon) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 && \text{on } \Sigma_{N,T}, \\ (\nabla u_\delta^\varepsilon + k_\delta(s_\delta^\varepsilon) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= -\frac{\xi^\varepsilon(u_\delta^\varepsilon)}{k_\delta(1)c} + \frac{w_\delta^\varepsilon}{c} + \frac{h^\varepsilon(u_\delta^\varepsilon) \psi^\varepsilon(w_\delta^\varepsilon)}{c} && \text{on } \Sigma_{in,T}, \\ s_\delta^\varepsilon(x, 0) &= s_{0,\delta}^\varepsilon && \text{in } \Omega, \\ \partial_t w_\delta^\varepsilon - \varepsilon \Delta w_\delta^\varepsilon &= r^\varepsilon + \frac{\xi^\varepsilon(u_\delta^\varepsilon)}{k_\delta(1)c} - \frac{w_\delta^\varepsilon}{c} - \frac{h^\varepsilon(u_\delta^\varepsilon) \psi^\varepsilon(w_\delta^\varepsilon)}{c} && \text{in } \Sigma_{in,T}, \\ \nabla w_\delta^\varepsilon \cdot \boldsymbol{\nu}_2 &= 0 && \text{on } \partial\Sigma_{in} \times [0, T], \\ w_\delta^\varepsilon(x, 0) &= w_{0,\delta}^\varepsilon && \text{in } \Sigma_{in}, \end{aligned} \tag{22}$$

where  $\nu_2$  is the outer normal of  $\partial\Sigma_{\text{in}}$  and the regularized functions fulfill the following assumptions.  $f^\varepsilon(s, \cdot)$  and  $f^\varepsilon(\cdot, x, t)$  shall be in  $C^2(\Omega_T)$  and  $C^2([0, 1])$  respectively,  $f^\varepsilon \rightarrow f$  uniformly on  $[0, 1] \times \Omega_T$  and  $\|f^\varepsilon\|_{C^0([0, 1] \times \Omega_T)} \leq \|f\|_{C^0([0, 1] \times \Omega_T)}$ .  $r^\varepsilon$  shall be in  $C^2(\Omega_T)$ ,  $\|r^\varepsilon\|_{L^\infty(\Omega_T)} \leq \|r\|_{L^\infty(\Omega_T)}$  and  $\|r^\varepsilon - r\|_{H^{1,2}(\Omega_T)} \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . Furthermore, we require that  $\xi^\varepsilon \in C^2(\mathbb{R})$  and that  $\xi^\varepsilon$  is nonnegative and nondecreasing such that  $\xi^\varepsilon(p_\delta^\varepsilon) = (p_\delta)_+$  if  $p_\delta \in \mathbb{R}^- \cup (\varepsilon, \infty)$  and  $|\xi^\varepsilon(p_\delta^\varepsilon)| \leq 1 + \varepsilon^{-1}$  and analog for  $u_\delta$ .  $\chi^\varepsilon \in C_0^\infty(\Sigma_{\text{out}})$  is a cut-off function with  $\chi^\varepsilon = 1$  if  $\text{dist}(x, \partial\Sigma_{\text{out}}) > \varepsilon$ . Additionally,  $h^\varepsilon$  shall be in  $C^2(\mathbb{R})$ ,  $h^\varepsilon \geq 0$ ,  $h^\varepsilon = 0$  if  $u_\delta^\varepsilon \geq 0$ ,  $h^\varepsilon \leq 0$  and  $h^\varepsilon \rightarrow h$  uniformly on  $\mathbb{R}$ . Moreover, we require  $\psi^\varepsilon$  to be in  $C^2(\mathbb{R})$ , to be nonnegative and that  $\psi^\varepsilon \rightarrow \psi$  uniformly on  $\mathbb{R}^+$ . Finally, the regularized initial values  $w_{0,\delta}^\varepsilon$  and  $s_{0,\delta}^\varepsilon$  shall be in  $C^2(\overline{\Sigma_{\text{in}}})$  or  $C^3(\overline{\Omega})$  respectively with  $\|w_{0,\delta}^\varepsilon - w_{0,\delta}\|_{L^1(\Sigma_{\text{in}})} \rightarrow 0$  and  $\|s_{0,\delta}^\varepsilon - s_{0,\delta}\|_{L^2(\Omega)} \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . They also have to fulfill suitable compatibility conditions.

**Lemma 3.14** (Existence of a smooth solution pair  $(u_\delta^\varepsilon, w_\delta^\varepsilon)$  of system (22)). *Suppose that the regularized functions in system (22) fulfill the just mentioned assumptions. Then there exists a solution pair  $(u_\delta^\varepsilon, w_\delta^\varepsilon)$  with  $u_\delta^\varepsilon \in C^0(\overline{\Omega_T})$ ,  $\nabla u_\delta^\varepsilon \in C^0(\overline{\Omega_T})$ ,  $w_\delta^\varepsilon \in C^0(\overline{\Sigma_{\text{in},T}})$ ,  $\Delta w_\delta^\varepsilon \in C^0(\overline{\Sigma_{\text{in},T}})$ ,  $\partial_t w_\delta^\varepsilon \in C^0(\overline{\Sigma_{\text{in},T}})$ . Furthermore, we have  $s_\delta^\varepsilon \in C^0(\overline{\Omega_T})$  and  $\nabla s_\delta^\varepsilon \in C^0(\overline{\Omega_T})$ .*

*Proof.* The regularized coefficient functions fulfill all necessary assumptions in the respective theorems in Section 5 of [30], which yield the required result.  $\square$

For the solution  $(u_\delta^\varepsilon, w_\delta^\varepsilon)$  we can now derive a maximum principle and obtain  $L^\infty$ -estimates independent of  $\varepsilon$  and  $\delta$ .

**Lemma 3.15** (Maximum principle for the approximating solution pair  $(u_\delta^\varepsilon, w_\delta^\varepsilon)$ ). *For the solution pair  $(u_\delta^\varepsilon, w_\delta^\varepsilon)$  of 22 we have the following inequalities*

$$\|w_\delta^\varepsilon\|_{L^\infty(\Sigma_{\text{in},T})} \leq C \text{ and } \|(u_\delta^\varepsilon)_+\|_{L^\infty(\Sigma_{\text{in},T})} \leq \max\{\varepsilon, c_1 \|w_\delta^\varepsilon\|_{L^\infty(\Sigma_{\text{in},T})} + c_2\},$$

where the constants  $C, c_1$  and  $c_2$  do not depend on  $\delta$  and  $\varepsilon$ .

*Proof.* The proof follows ideas presented in [24]. We omit the indices  $\delta$  and  $\varepsilon$  in this proof. Let  $u_s \in C^2(\overline{\Omega})$  be the solution of

$$\begin{aligned} \Delta u_s + \nabla k(\Phi^{-1}(u_s)) \cdot \mathbf{e} + F &= 0 && \text{in } \Omega, \\ u_s &= 0 && \text{on } \Sigma_{\text{in}}, \\ (\nabla u_s + k(\Phi^{-1}(u_s)) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 && \text{on } \Sigma_N, \\ (\nabla u_s + k(\Phi^{-1}(u_s)) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= -\delta^{-1} \xi^\varepsilon(u_s) \chi^\varepsilon && \text{on } \Sigma_{\text{out}}, \end{aligned}$$

where  $F = \|f(x, t, s)\|_{C^0(\Omega_T \times [0, 1])}$ . Analogous to [24] one can show that  $u_s$  is nonnegative. Therefore the term  $\nabla k(\Phi^{-1}(u_s))$  is unproblematic as  $\Phi_\delta^{-1}(u_s^\delta)$  tends to zero as  $\delta \rightarrow 0$  for  $u_s^\delta \geq 0$ . Thus we can bound  $\Delta u_s + \nabla k(\Phi^{-1}(u_s)) \cdot \mathbf{e}$  and see that hence  $u_s^\delta \rightarrow 0$  as  $\delta \rightarrow 0$  on  $\Sigma_{\text{out}}$ . A classical maximum principle yields that  $\|u_s\|_{C^1(\overline{\Omega})} \leq C$ , with  $C$  independent of  $\delta$  and  $\varepsilon$ .

We fix  $t \in (0, T)$ , define  $m_t := \max_{(x, \tau) \in \overline{\Sigma_{\text{in},T}}} u_+(x, \tau)$  and set  $U_s := u_s(x) + m_t$  for  $x \in \overline{\Omega}$ . With the same techniques as in [1] one can prove that

$$u(x, \tau) \leq U_s(x) \quad \forall (x, \tau) \in \overline{\Omega_t}. \tag{23}$$

We have that

$$u(x(\hat{t}), \hat{t}) = U_s(x(\hat{t})) \text{ for some } (x(\hat{t}), \hat{t}) \in \overline{\Sigma_{\text{in}}} \times [0, t] \quad (24)$$

with  $m_t = u_+(x(\hat{t}), \hat{t})$ . We suppose that  $m_t > 0$ , otherwise the proof is easy. Then (23) and (24) yield  $\nabla U_s(x(\hat{t}), \hat{t}) \cdot \boldsymbol{\nu} \leq \nabla u(x(\hat{t}), \hat{t}) \cdot \boldsymbol{\nu}$ , which in turn leads to

$$\frac{\xi^\varepsilon(m_t)}{k(1)c} \leq \frac{w(x(\hat{t}), \hat{t})}{c} - (\nabla u_s(x(\hat{t}))) + k(\Phi^{-1}(u_s(x(\hat{t}))) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} \leq \frac{v(t)}{c} + c_2, \quad (25)$$

where  $c_2$  does not depend on  $\varepsilon$  and  $\delta$  and

$$v(t) := \max_{\tau \in [0, t]} W(\tau), \quad W(\tau) := \max_{x \in \Sigma_{\text{in}}} w(x, \tau).$$

Then  $v$  is nonnegative, nondecreasing and Lipschitz continuous on  $[0, T]$ . For  $W(t) = w(x(t), t)$  with  $x(t)$  in  $\Sigma_{\text{in}}$  we have  $\Delta w(x(t), t) \leq 0$ . At the points where  $W(t)$  is differentiable, there holds

$$W'(\tau) \leq -\frac{W(\tau)}{c} + \|r\|_{L^\infty(\Sigma_{\text{in}, T})} + \frac{\xi^\varepsilon(u(x(\tau), \tau))}{k(1)c} \leq \frac{\xi^\varepsilon(m_t)}{k(1)c} - \frac{W(\tau)}{c} + C. \quad (26)$$

At first we assume that  $v(t) = W(\tau)$  for some  $0 \leq \tau < t$ . Since then  $v(s) = W(\tau)$  for all  $s \in [\tau, t]$ , we have  $v' = 0$  on  $(\tau, t)$ . Thus let  $v(t) = W(t)$  for some  $t \in (0, T)$ . Then (25) and (26) lead to  $v'(t) \leq C$  and hence  $v(t) \leq C$  for any  $t \in [0, T]$  and a constant  $C$  not depending on  $\delta$  and  $\varepsilon$ . Therefore we have

$$\|w_\delta^\varepsilon\|_{L^\infty(\Sigma_{\text{in}, T})} \leq C \quad (27)$$

and with (25) and (27) also

$$\|(u_\delta^\varepsilon)_+\|_{L^\infty(\Sigma_{\text{in}, T})} \leq \max\{\varepsilon, c_1\|w_\delta^\varepsilon\|_{L^\infty(\Sigma_{\text{in}, T})} + c_2\}.$$

□

If we now pass to the limit in  $\varepsilon$  in (22) and obtain the following estimates.

**Lemma 3.16** ( $L^\infty$ -estimates for  $u_\delta$  and  $w_\delta$ ). *For the solution pair  $(u_\delta, w_\delta)$  of (13) we have the following inequalities*

$$\|u_\delta\|_{L^\infty(\Omega_T)} + \|w_\delta\|_{L^\infty(\Sigma_{\text{in}, T})} \leq C, \quad (28)$$

$$\|(u_\delta)_+\|_{L^\infty(\Sigma_{\text{in}, T})} \leq c_1\|w_\delta\|_{L^\infty(\Sigma_{\text{in}, T})} + c_2, \quad (29)$$

where the constants  $C, c_1$  and  $c_2$  do not depend on  $\delta$ .

*Proof.* First let us assume that  $s_{0,\delta} \in L^\infty(\Omega) \cap H^{1,2}(\Omega)$ ,  $w_{0,\delta} \in L^\infty(\Sigma_{\text{in}}) \cap H^{1,2}(\Sigma_{\text{in}})$  and that they fulfill suitable compatibility conditions. Approximating  $s_{0,\delta}$  and  $w_{0,\delta}$  by the initial conditions of (22) with  $s_{0,\delta}^\varepsilon \rightarrow s_{0,\delta}$  in  $H^{1,2}(\Omega)$  and  $w_{0,\delta}^\varepsilon \rightarrow w_{0,\delta}$  in  $H^{1,2}(\Sigma_{\text{in}})$ , we obtain for the solution pair  $(u_\delta^\varepsilon, w_\delta^\varepsilon)$  of (22) the following a priori estimates

$$\|u_\delta^\varepsilon\|_{L^\infty(\Omega_T)} \leq C, \quad p_\delta^\varepsilon \leq p_{\max} \text{ on } \Omega_T, \quad (30)$$

$$\begin{aligned} & \|\nabla u_\delta^\varepsilon\|_{L^2(\Omega_T)} + \|\partial_t u_\delta^\varepsilon\|_{L^2(\Omega_T)} + \|\partial_t s_\delta^\varepsilon\|_{L^2(\Omega_T)} \\ & + \|\text{div}(\nabla u_\delta^\varepsilon + k_\delta(s_\delta^\varepsilon) \cdot \mathbf{e})\|_{L^2(\Omega_T)} + \|\partial_t w_\delta^\varepsilon\|_{L^2(\Sigma_{\text{in}, T})} \leq C(\delta), \end{aligned} \quad (31)$$

$$\begin{aligned} & \max_{t \in [0, T]} \|\partial_t w_\delta^\varepsilon(\cdot, t)\|_{L^1(\Sigma_{\text{in}})} + \max_{t \in [0, T]} \|\nabla w_\delta^\varepsilon(\cdot, t)\|_{L^1(\Sigma_{\text{in}})} \\ & \leq C(\delta)(1 + \|\partial_t w_\delta^\varepsilon(\cdot, 0)\|_{L^1(\Sigma_{\text{in}})} + \|\nabla w_\delta^\varepsilon(\cdot, 0)\|_{L^1(\Sigma_{\text{in}})}), \end{aligned} \quad (32)$$

where all constants  $C$  and  $p_{max}$  are independent of  $\varepsilon$  and the constants in (30) do not depend on  $\delta$  either. Here Lemma 3 in [37] yields (30) and the estimates in (31) can be obtained by a testing procedure. Applying the vanishing viscosity method (see for instance [4]) results in (32). We can extract subsequences  $u_\delta^\varepsilon \rightarrow u_\delta$  in  $L^2(\Omega_T)$  with  $\nabla u_\delta^\varepsilon \rightarrow \nabla u_\delta$  in  $L^2(\Omega_T)$  and  $w_\delta^\varepsilon(\cdot, t) \rightarrow w_\delta(\cdot, t)$  in  $L^1(\Sigma_{in})$  with  $\partial_t w_\delta^\varepsilon(\cdot, t) \rightarrow \partial_t w_\delta(\cdot, t)$  in  $L^1(\Sigma_{in})$  for  $t \in [0, T]$  as  $\varepsilon \rightarrow 0$ , where  $(u_\delta, w_\delta)$  is the solution pair of (13). To verify that the limit of  $w_\delta^\varepsilon$  fulfills the ODE in (13), we test with a function in  $C_0^\infty(\Sigma_{in})$  and apply the fundamental lemma of calculus of variations to obtain that  $\partial_t w_\delta - r_\delta - (u_\delta)_+/(k_\delta(1)c) + w_\delta/c + (h_\delta(u_\delta)\psi(w_\delta))/c = 0$  a.e. on  $\Sigma_{in}$  for all  $t \in [0, T]$ . Thus (28) and (29) follow. Let us now suppose that  $(s_{0,\delta}, w_{0,\delta})$  are only in  $L^\infty(\Omega)$  or  $L^\infty(\Sigma_{in})$  respectively. We approximate  $(s_{0,\delta}, w_{0,\delta})$  by  $C_0^\infty$ -functions  $(s_{0,\delta}^\eta, w_{0,\delta}^\eta)$  such that  $s_{0,\delta}^\eta \rightarrow s_{0,\delta}$  in  $L^1(\Omega)$  and  $w_{0,\delta}^\eta \rightarrow w_{0,\delta}$  in  $L^1(\Sigma_{in})$  as  $\eta \rightarrow 0$  and denote the corresponding solution pair of the system (13)  $(s_\delta^\eta, w_\delta^\eta)$ . Estimates (17) and (19) imply that  $w_\delta^\eta \rightarrow w_\delta$  in  $C^0([0, T]; L^1(\Sigma_{in}))$  and  $s_\delta^\eta(\cdot, t) \rightarrow s_\delta(\cdot, t)$  in  $L^1(\Omega)$  for a.e.  $t \in [0, T]$  as  $\eta \rightarrow 0$ . Moreover, (17) renders  $\int_{\Sigma_{in}} |(u_\delta^\eta)_+ - (u_\delta)_+| + \psi(w_\delta)|h_\delta(u_\delta^\eta) - h_\delta(u_\delta)| \rightarrow 0$  as  $\eta \rightarrow 0$ . Thus a priori estimates (28) and (29), which hold true for  $(u_\delta^\eta, w_\delta^\eta)$  also hold true for  $(u_\delta, w_\delta)$ .  $\square$

**Remark 3.17.** Note that the parabolic maximum principle yields that  $s_\delta \in [0, 1]$  a.e. in  $\Omega_T$  (cf. [37] for details).

**Lemma 3.18.** There exists  $C < \infty$  independent of  $\delta$  such that, for all  $\delta > 0$ ,

$$\int_{\Omega_T} k_\delta(s_\delta) |\nabla s_\delta|^2 \leq C \quad \text{and} \quad u \in L^2((0, T), H^{1,2}(\Omega)). \quad (33)$$

*Proof.* To prove (33), one first proves for a constant  $C$ , independent of  $\delta$ , the estimate

$$\int_{\Omega_T} k_\delta(s_\delta) |\nabla p_\delta|^2 \chi_{\{s_\delta \geq 1/2\}} + k_\delta(s_\delta) \partial_s p_\delta(s_\delta) |\nabla s_\delta|^2 \chi_{\{s_\delta < 1/2\}} \leq C,$$

where  $\chi_D$  denotes the characteristic function of a set  $D$ . To this end, we test the weak formulation (15) with suitable test functions. For  $s_\delta < 1/2$  we choose  $\eta_-(s_\delta) = (s_\delta - 1/2)_- + 1/2$  and for  $s_\delta \geq 1/2$  we pick  $\eta_+(p_\delta) = (p_\delta - \bar{p})_+$ , where  $p_\delta(0) < \bar{p} \leq p_\delta(1/2)$ . Theorem 3.13 and Remark 3.17 guarantee the boundedness of  $\eta_-(s_\delta)$  and  $\eta_+(p_\delta)$ . For more details we refer to [37, 38].  $\square$

### 3.4 Passing to the limit in $\delta$

**Theorem 3.19.** Let the assumptions of Section 2 and Subsection 3.1 be fulfilled and let  $(u_\delta, w_\delta)$  be the unique solution pair of the regularized system (13). Then for a subsequence  $\delta \rightarrow 0$  there holds

$$s_\delta \xrightarrow{*} s \quad \text{in } L^\infty(\Omega_T), \quad (34)$$

$$u_\delta \rightharpoonup u \quad \text{in } L^2([0, T], H^{1,2}(\Omega)), \quad (35)$$

$$w_\delta \rightarrow w \quad \text{in } C^0([0, T], L^1(\Sigma_{in})),$$

for appropriate limit functions  $s : \Omega_T \rightarrow [0, 1]$ ,  $u : \Omega_T \rightarrow \mathbb{R}$  and  $w : \Sigma_{in,T} \rightarrow \mathbb{R}$ . The limits satisfy

$$\begin{aligned}
u &\in \Phi(s) \quad \text{in } \Omega_T, \quad \partial_t s = \operatorname{div}(\nabla u + k(s) \cdot \mathbf{e}) + f(s) \quad && \text{in } \mathcal{D}'(\Omega_T), \\
-(\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &\geq 0 \quad && \text{on } \overline{\Sigma}_{out,T} \cup \Sigma_{N,T}, \\
u &\leq \tilde{\Phi}(1) \quad \text{and } k^2(s)s - k^2(1)1 \leq 0 \quad && \text{on } \Sigma_{out,T}, \\
0 &\leq ((\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu}) \cdot (k^2(s)s - k^2(1)1) \quad && \text{on } \Sigma_{out,T}, \\
&&& (36) \\
(\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 \quad && \text{on } \Sigma_{N,T}, \\
(\nabla u + k(s) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= g(u, w) \quad && \text{on } \Sigma_{in,T}, \\
s &= s_0 \quad && \text{in } \Omega_0, \\
\partial_t w &= (r - g(u, w)) \quad && \text{in } \Sigma_{in,T}, \\
w &= w_0 \quad && \text{on } \Sigma_{in,0},
\end{aligned}$$

where

$$\begin{aligned}
g(u, w) &= -\frac{u_+}{k(1)c} + \frac{w}{c} + \frac{h(u)\psi(w)}{c}, \\
h(u) &= \int_{u_-}^0 \frac{1}{k(\Phi^{-1}(\tau))} d\tau, \quad \psi(w) = \min \left\{ 1, \left( \frac{w}{\sigma} \right)_+ \right\}.
\end{aligned}$$

The traces in (36)<sub>3</sub>–(36)<sub>7</sub> exist in the sense of distributions. Furthermore, there holds  $w \geq 0$  a.e. on  $\Sigma_{in}$  for all  $t \in [0, T]$ .

*Proof.* The assumptions in Section 2 and Subsection 3.1 and the a priori bounds of Subsection 3.3 yield (34), (35),  $s : \Omega_T \rightarrow [0, 1]$  and  $f_\delta \rightharpoonup f$  in  $L^2(\Omega_T)$ . Affinity of  $f(x, t, \cdot)$  on  $(0, \lambda)$  for small  $\lambda$  and the compactness of  $s_\delta$  on  $(\lambda, 1]$  results in  $f_\delta \rightarrow f$  in  $L^2(\Omega_T)$ . Using again Remark 3.17, the assumption  $k(0) = 0$  and the compactness of  $s_\delta$  on  $(\lambda, 1]$  provides  $k_\delta(s_\delta) \rightarrow k(s)$  in  $L^2(\Omega_T)$  and  $\nabla k_\delta(s_\delta) \rightharpoonup \nabla k(s)$  in  $(L^2(\Omega_T))^3$  for  $\delta \rightarrow 0$ . Therefore, we can identify (36)<sub>1</sub> as the distributional limit of (13)<sub>1</sub>.

**Compactness of the families  $s_\delta$  and  $u_\delta$ .** We outline the ideas of the proof and refer to [37] for more details. As  $\Phi'_\delta$  is unbounded, we only aim at compactness for  $u_\delta$  away from regions with maximal saturation. To prove compactness for the family  $u_\delta$ , we use a sequence  $\varepsilon \rightarrow 0$  and an associated sequence of cut-off functions  $\alpha_\varepsilon \in C^\infty(\Omega, [0, 1])$  with  $\alpha_\varepsilon(x) = 1$  for all  $x \in \Omega$  with  $\operatorname{dist}(x, \partial\Omega) \geq \varepsilon$ . If we consider  $\eta_\varepsilon(\xi) := (\xi + \varepsilon)_-$  for fixed  $\varepsilon$ , it can be shown that the family  $\alpha_\varepsilon \cdot \eta_\varepsilon(u_\delta)$  is compact in  $L^2(\Omega_{T-\varepsilon})$  and that for a subsequence  $\delta \rightarrow 0$  there holds

$$\eta_\varepsilon(u_\delta) \xrightarrow{\delta \rightarrow 0} \eta_\varepsilon(u) \quad \text{in } L^2(\Omega_T) \quad \text{for any } \varepsilon > 0. \quad (37)$$

Here, the compactness of the family  $\alpha_\varepsilon \cdot \eta_\varepsilon(u_\delta)$  in  $L^2(\Omega_{T-\varepsilon})$  can be deduced using the Riesz characterization of compact sets. The compactness of the family  $s_\delta$  away from regions with zero saturation follows similarly.

For the proof of relation (36)<sub>2</sub>, the boundary conditions on the Neumann and the outflow boundary and the initial conditions we refer to [37]. We only remark

that (36)<sub>2–4</sub>, (36)<sub>6</sub> and (36)<sub>8</sub> follow more or less straightforward, while the proof of (36)<sub>5</sub> is much more challenging and uses defect measures and a compensated compactness argument. We now turn to the proof of (36)<sub>7</sub>. The problem we have to cope with on the infiltration boundary is the fact that  $h(u)$  is not Lebesgue-integrable for  $S := \{(x, t) \in \Omega_T \mid s(x, t) = 0\}$  not being a null set in the sense of Lebesgue. Since we are not able to bound  $h_\delta(u_\delta)$  on  $S$  independently of  $\delta$ ,  $h_\delta(u_\delta)$  is not necessarily Lebesgue-integrable on  $S$  either. To prove convergence on the infiltration boundary, we exploit that  $p_-(x, t) = -h(u(x, t))$  and that the water pressure  $p(x, t)$  equals the capillary pressure  $p_c(s)$ . Thanks to Assumption 2.1, we know that  $(-p)_+$  is bounded and therefore that  $h(u)$  is bounded, too.

Before analyzing the limiting process for  $\delta \rightarrow 0$ , we show that our boundary condition on  $\Sigma_{\text{in}}$

$$(\nabla u_\delta + k_\delta(s_\delta) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} = -\frac{(u_\delta)_+}{k_\delta(1)c} + \frac{w_\delta}{c} + \frac{h_\delta(u_\delta)\psi(w_\delta)}{c}$$

is well-defined in the sense of traces. Since  $w_\delta$  is only defined on  $\Sigma_{\text{in}}$ , we do not have to discuss the terms solely depending on  $w_\delta$ . Due to Theorem 3.12 c), the function  $(u_\delta)_+$  has a well-defined trace on  $\Sigma_{\text{in},T}$ , which leaves us the term  $h_\delta(u_\delta)$  and, particularly, because of Theorem 3.12 e),

$$h'_\delta(u_\delta) = \frac{(1 - \text{sign}(u_\delta))}{2(k_\delta(\Phi_\delta^{-1}(1/2(u_\delta - |u_\delta|))))}.$$

For simplicity, we restrict our analysis to the areas where  $u_\delta \leq 0$ . This is not a constraint, because  $h'_\delta(u_\delta)$  is bounded in the complement of these areas in  $\Omega_T$  anyway. We arrive at  $h'_\delta(u_\delta) \sim 1/(k_\delta(\Phi_\delta^{-1}(u_\delta)))$ . Obviously,  $h'_\delta(u_\delta)$  is unbounded if  $s_\delta = \Phi_\delta^{-1}(u_\delta) \rightarrow 0$ . As already discussed at the beginning of this paragraph, this would be a contradiction to the fact that  $(-p)_+$  is bounded due to assumption (11). This proves that  $h'_\delta(u_\delta)$  is in fact bounded and, therefore, that  $h_\delta(u_\delta)$  has a well-defined trace.

Since  $\nabla h_\delta(\eta_\varepsilon(u_\delta))$  is bounded in  $(L^2(\Omega_T))^3$ , we get a subsequence, converging weakly to a limit  $\tilde{h}$ . Using (37) and Assumption 2.1, one can then easily prove that  $\tilde{h} = \nabla h(\eta_\varepsilon(u))$  and  $h_\delta(\eta_\varepsilon(u_\delta)) \rightarrow h(\eta_\varepsilon(u))$  in  $L^2(\Omega_T)$ . If we now account for the fact that  $h_\delta(u_\delta)$  only depends on the negative part of  $u_\delta$ , we get that  $\nabla h_\delta(u_\delta) \rightharpoonup \nabla h(u)$  and  $h_\delta(u_\delta) \rightarrow h(u)$  in  $(L^2(\Omega_T))^3$  and  $L^2(\Omega_T)$ , respectively, for  $\delta \rightarrow 0$ . This results in  $h_\delta(u_\delta) \rightarrow h(u)$  in  $L^2(\Sigma_{\text{in},T})$ . The convergence  $u_\delta \rightarrow u$  in  $L^2(\Sigma_{\text{in},T})$  implies directly the convergence  $(u_\delta)_+ \rightarrow u_+$  in  $L^2(\Sigma_{\text{in},T})$  for  $\delta \rightarrow 0$ . Applying these results and Gronwall's Lemma, one gets  $w_\delta \rightarrow w$  in  $C^0([0, T], L^1(\Sigma_{\text{in}}))$ , implying that  $w_x(0) = w_0$  and  $w \geq 0$  a.e. on  $\Sigma_{\text{in}}$  for all  $t \in [0, T]$ .  $\square$

## 4 Discretization and convex minimization

In this section we introduce an implicit-explicit time discretization of System (9), followed by a space discretization with finite elements for the Richards equation. We are bound to this particular discretization in the sense that it allows for spatial convex minimization problems in the porous medium that on the discrete level can be treated by monotone multigrid. We refer to [13] for a detailed analytical and numerical study of this approach. Here, we extend

the approach in [13], replacing Dirichlet conditions by nonlinear physical Robin conditions. For the treatment of Robin conditions, in which the physical pressure appears without nonlinear deformation, see also [12]. The ODE's for the surface water are treated explicitly. For another example of coupling Richards' equation with a single ODE representing the water height of a lake given by a compartment model, we refer to [10, Sec. 4.3]. See [8] and [14] for implicit treatments of the coupling of Richards' equation with the shallow water equations, where in the latter article—as in this paper—the coupling is also given by a Robin transmission condition.

## 4.1 Time discretization

We introduce our special implicit-explicit time discretization of System (9), first in a strong and then in a weak formulation. Special attention will be given to the nonlinear Robin condition on  $\Sigma_{\text{in}}$ . The Signorini-type condition will lead to a variational inequality.

### 4.1.1 Strong formulation

Given  $s^n$  in  $\Omega$  and  $w^n$  as well as  $r^{n+1}$  on  $\Sigma_{\text{in}}$  for time steps  $t_n$  and  $t_{n+1}$ , respectively, with  $\tau = t_{n+1} - t_n$ , we define  $(s^{n+1}, w^{n+1})$  at time step  $t_{n+1}$  as follows.

$$\begin{aligned}
u^{n+1} &\in \Phi(s^{n+1}), \quad s^{n+1} = \tau \operatorname{div}(\nabla u^{n+1} + k(s^n) \cdot \mathbf{e}) + \tau f(s^{n+1}) + s^n \quad \text{in } \Omega, \\
u^{n+1} &\leq 0 \quad \text{and} \quad -(\nabla u^{n+1} + k(s^n) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} \geq 0 \quad \text{on } \Sigma_{\text{out}}, \\
0 &= ((\nabla u^{n+1} + k(s^n) \cdot \mathbf{e}) \cdot \boldsymbol{\nu}) \cdot u^{n+1} \quad \text{on } \Sigma_{\text{out}}, \\
&\quad (38) \\
(\nabla u^{n+1} + k(s^n) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= 0 \quad \text{on } \Sigma_N, \\
(\nabla u^{n+1} + k(s^n) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} &= g(u^{n+1}, w^n) \quad \text{on } \Sigma_{\text{in}}, \\
w^{n+1} &= \tau(r^{n+1} - g(u^{n+1}, w^n)) + w^n \quad \text{in } \Sigma_{\text{in}}.
\end{aligned}$$

Here, boundary condition (38)<sub>3</sub>, (38)<sub>4</sub> is the discretization of the original Signorini-type condition (2) which is equivalent to (9)<sub>3</sub>–(9)<sub>5</sub>. Boundary condition (38)<sub>6</sub> needs special consideration. First, we rewrite  $g$  on  $\Sigma_{\text{in}} \times \Sigma_{\text{in}}$  (defined in (9)) by inserting the inverse Kirchhoff transformation  $\kappa^{-1} : u \mapsto p$  given in Equations (8) and obtain

$$\begin{aligned}
-g(u^{n+1}, w^n) + \frac{w^n}{c} &= \frac{u_+^{n+1}}{k(1)c} - \frac{h(u^{n+1}) \min\{1, (\frac{w^n}{\sigma})_+\}}{c} \\
&= c^{-1} \kappa^{-1}(u^{n+1})_+ + c^{-1} \kappa^{-1}(u^{n+1})_- \cdot \psi(w^n).
\end{aligned}$$

Therefore, with  $p^{n+1} := \kappa^{-1}(u^{n+1})$ , boundary condition (38)<sub>6</sub> reads

$$\begin{aligned}
(k(s^{n+1}) \nabla p^{n+1} + k(s^n) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} + c^{-1} p_+^{n+1} + c^{-1} p_-^{n+1} \cdot \psi(w^n) = \\
(\nabla u^{n+1} + k(s^n) \cdot \mathbf{e}) \cdot \boldsymbol{\nu} + c^{-1} \kappa^{-1}(u^{n+1})_+ + c^{-1} \kappa^{-1}(u^{n+1})_- \cdot \psi(w^n) = c^{-1} w^n
\end{aligned}$$

on  $\Sigma_{\text{in}}$  and is a nonlinear Robin condition both in  $p^{n+1}$  and in  $u^{n+1}$ . The special structure of it when formulated in  $u^{n+1}$  stems from the fact that the real functions

$$\kappa_\xi^* : u \mapsto c^{-1} \kappa^{-1}(u)_+ + c^{-1} \kappa^{-1}(u)_- \cdot \psi(w^n(\xi)), \quad (39)$$

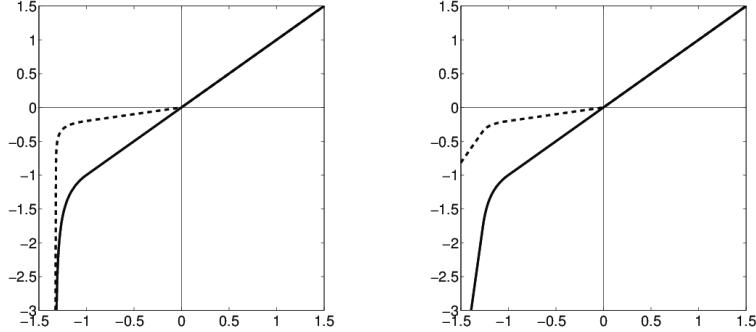


Figure 2: Left: degenerate  $\kappa_\xi^* = \kappa^{-1}$  with  $\Phi^{-1}(0) = -1.3245 \cdot |p_b| > -\infty$  by Brooks–Corey, (bold line, data from Section 5) and with weight  $\psi(w^n(\xi)) = 0.2$  for  $u_-$  (dotted line). Right: regularized  $(\kappa_\xi^*)_\delta$  with  $\delta^2 = 0.1$  and  $\Phi^{-1}(0) = -\infty$ , (bold line, construction by  $k_\delta(s) := \max(k(s), \delta^2)$  as in [10, Sec. 1.4.3]) and with weight  $\psi(w^n(\xi)) = 0.2$  for  $u_-$  (dotted line).

which, for  $\xi \in \Sigma_{\text{in}}$  almost everywhere, induce a superposition operator on  $\Sigma_{\text{in}}$ , are increasing, cf. [11]. This observation is essential for our way to determine  $u^{n+1}$  that we present below. In fact, for  $w^n(\xi) \geq \sigma$ , the function  $c \cdot \kappa_\xi^*$  is nothing but the real function  $\kappa^{-1}$  that induces the inverse Kirchhoff transformation (8). Otherwise, the factor  $\psi(w^n(\xi)) = (\frac{w^n(\xi)}{\sigma})_+ < 1$  just leads to a smaller weight for negative values of  $p$ , see Fig. 2. Finally, note that system (38) only contains an explicit coupling of ground and surface water in the sense that the solution  $u^{n+1}$  in the porous medium depends implicitly on the known value  $w^n$  from the previous time step while  $w^{n+1}$  can be explicitly determined once  $u^{n+1}$  is known.

#### 4.1.2 Weak formulation

In contrast to the considerations in Section 3, our weak formulation of the spatial problem (38) is based on the second line of (7), concentrating on  $u^{n+1}$  instead of  $s^{n+1}$  as the unknown on  $\Omega$ . Recall that by the Kirchhoff transformation (6), we have  $u \geq \Phi(0)$ . Depending on the kind of degeneracy of  $k(p_c^{-1}(q))$  for  $q \rightarrow -\infty$ , we might have  $\Phi(0) = -\infty$  or  $-\infty < \Phi(0) < 0$  as in the case of Brooks–Corey parameter functions (see (51) and (52) or [10, Sec. 1.3] and Fig. 2). Anyway, for a weak formulation we should expect  $u^{n+1}$  to be an element of the convex subset

$$\mathcal{K} := \{v \in H^1(\Omega) : v \geq \Phi(0), v|_{\Sigma_{\text{out}}} \leq 0\}$$

of the Sobolev space  $H^1(\Omega)$ . Together with the Signorini-type outflow condition (2), which is also encoded in  $\mathcal{K}$ , these constraints lead to the following weak interpretation of (38)<sub>1</sub>–(38)<sub>6</sub> in terms of a variational inequality for  $u^{n+1} \in \mathcal{K}$

(for details consult [10, Sec. 1.5.2/3, 3.4.5], where  $\mathbf{e} = -e_z$  :

$$\begin{aligned}
& \int_{\Omega} (\Phi^{-1}(u^{n+1}) - \tau f(\Phi^{-1}(u^{n+1}))) \cdot (v - u^{n+1}) + \tau \int_{\Sigma_{\text{in}}} \kappa_{\xi}^*(u^{n+1}) \cdot (v - u^{n+1}) \\
& + \tau \int_{\Omega} \nabla u^{n+1} \cdot \nabla (v - u^{n+1}) - \int_{\Omega} \Phi^{-1}(u^n) \cdot (v - u^{n+1}) \\
& + \tau \int_{\Omega} k(\Phi^{-1}(u^n)) \mathbf{e} \cdot \nabla (v - u^{n+1}) - \tau \int_{\Sigma_{\text{in}}} \frac{w^n}{c} \cdot (v - u^{n+1}) \geq 0 \quad \forall v \in \mathcal{K}.
\end{aligned} \tag{40}$$

## 4.2 Convex theory

In this section, we show that under reasonable assumptions, the variational inequality (40) is equivalent to a uniquely solvable convex minimization problem.

Let  $\Psi_x$ ,  $x \in \Omega$ , and  $\Psi_{\xi}$ ,  $\xi \in \Sigma_{\text{in}}$ , be primitives of  $\Phi^{-1} - \tau f(x, t_{n+1}(\Phi^{-1}(\cdot)))$ ,  $x \in \Omega$ , and  $\tau \kappa_{\xi}^*$ ,  $\xi \in \Sigma_{\text{in}}$ , respectively, on  $[\Phi(0), \infty) \cap \mathbb{R}$ . Since  $\Phi^{-1} - \tau f \Phi^{-1}$  and  $\kappa_{\xi}^*$  are increasing (cf. Assumption 2.2),  $\Psi_x$  and  $\Psi_{\xi}$  are convex so that the functional

$$\begin{aligned}
F : v \mapsto & \int_{\Omega} \Psi_x(v) + \int_{\Sigma_{\text{in}}} \Psi_{\xi}(v) + \frac{\tau}{2} \int_{\Omega} |\nabla v|^2 \\
& - \int_{\Omega} \Phi^{-1}(u^n) \cdot v + \tau \int_{\Omega} k(\Phi^{-1}(u^n)) \mathbf{e} \cdot \nabla v - \tau \int_{\Sigma_{\text{in}}} \frac{w^n}{c} \cdot v \quad \forall v \in \mathcal{K}
\end{aligned} \tag{41}$$

is strictly convex. Due to the assumptions on  $k$ ,  $p_c$  and  $f$ , the functions  $\Psi_x$ ,  $x \in \Omega$ , are bounded and Lipschitz continuous. Note, however, that we have  $\kappa_{\xi}^*(u) \rightarrow -\infty$  for  $u \searrow \Phi(0)$  in case of  $w^n(\xi) > 0$  so that as least for  $\Phi(0) > -\infty$ , the transformation induced by  $\kappa_{\xi}^*$ ,  $\xi \in \Sigma_{\text{in}}$ , is ill-posed in general and the second integral in (40) is not well-defined as an  $L^1$ -integral, cf. left graphics in Fig. 2. However, if we regularize  $k$  and  $p_c$  according to Assumption 3.2, the regularized functions  $(\kappa_{\xi}^*)_{\delta}$ ,  $\xi \in \Sigma_{\text{in}}$ , turn out to be Lipschitz continuous on  $\mathbb{R}$  with Lipschitz constant  $\max_u (\kappa_{\xi}^*)'_{\delta}(u) = \frac{1}{c\delta^2} \min\{1, (\frac{w^n(\xi)}{\sigma})_+\}$  and the corresponding integral in (40) is well-defined, cf. right graphics in Fig. 2. The following theorem states unique solvability of the variational inequality (40) for regularized functions  $(\kappa_{\xi}^*)_{\delta}$ ,  $\xi \in \Sigma_{\text{in}}$ , if their inverses have uniformly bounded Lipschitz constants for  $\xi \in \Sigma_{\text{in}}$  or, equivalently, if the weighting factor  $\psi(w^n(\xi))$  is uniformly bounded away from zero on a submanifold of  $\Sigma_{\text{in}}$  with positive Hausdorff measure. One can even prove well-posedness of (40), cf. [10, Prop. 2.4.11].

**Theorem 4.1.** *Let  $\Psi_x$ ,  $x \in \Omega$ , be induced by coefficient functions according to Assumptions 2.1 or 3.2. Let  $\Psi_{\xi}$ ,  $\xi \in \Sigma_{\text{in}}$ , be induced by regularized coefficient functions according to Assumption 3.2 and, accordingly, let  $\kappa_{\xi}^*$ ,  $\xi \in \Sigma_{\text{in}}$ , be replaced by  $(\kappa_{\xi}^*)_{\delta}$  in (40). Furthermore, let  $w^n \in L^2(\Sigma_{\text{in}})$ . Then the variational inequality (40) is equivalent to the convex minimization problem*

$$u^{n+1} \in \mathcal{K} : \quad F(u^{n+1}) \leq F(v) \quad \forall v \in \mathcal{K}. \tag{42}$$

If  $w^n > 0$  holds on a nonempty submanifold of  $\Sigma_{\text{in}}$ , then (42) is uniquely solvable.

*Proof.* Both the equivalence proof and the proof for the unique solvability is a combination of the convex theory on the time-discretized Richards equation with Signorini-type boundary conditions presented in [10, Sec. 2.3] and [13] and with Robin conditions which can be found in [10, Sec. 3.4.1] and [12]. Here, Lipschitz continuity of the regularized  $(\kappa_\xi^*)_\delta$  is needed. The additional dependency of  $\Psi_x$  on  $x \in \Omega$  and  $\Psi_\xi$  on  $\xi \in \Sigma_{\text{in}}$  can be easily included in the theory, taking into account  $f(\cdot, t_{n+1}, s^{n+1}) \in L^\infty(\Omega)$  and  $\psi(w^n) \in L^2(\Sigma_{\text{in}})$  which, in particular, leads to well-definedness of all considered integrals.

The only issue we need to consider here is the lack of a Dirichlet boundary with positive Hausdorff measure which guarantees coercivity of the functional  $F$ . In our case, the latter is provided by the nonlinear Robin boundary condition with the help of Poincaré's inequality [19, p. 127] in  $H^1(\Omega)$  which gives

$$\left\| v - \frac{1}{|\Omega|} \int_\Omega v \right\|_{L^2(\Omega)} \leq c_\Omega \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in H^1(\Omega) \quad (43)$$

for some constant  $c_\Omega > 0$  depending on  $\Omega$ . In detail, for the nonempty submanifold  $\Sigma_w$  of  $\Sigma_{\text{in}}$  where  $w^n > 0$  holds, i.e.,  $w^n \geq w_0$  for a  $w_0 > 0$ , we prove that

$$\|v\|_{\Sigma_w} := \|\nabla v\|_{L^2(\Omega)} + \|v\|_{L^2(\Sigma_w)} \quad \forall v \in H^1(\Omega) \quad (44)$$

defines a norm on  $H^1(\Omega)$  which is equivalent to  $\|\cdot\|_{H^1(\Omega)}$ . To see this, note that by the trace theorem [15, p. 1.61], there is a  $c_{\Sigma_w} > 0$  such that the estimate

$$\|v\|_{\Sigma_w} \leq c_{\Sigma_w} \|v\|_{H^1(\Omega)} \quad \forall v \in H^1(\Omega)$$

holds for the seminorm  $\|\cdot\|_{\Sigma_w}$ . Moreover, if  $\|v\|_{\Sigma_w} = 0$ , then we have  $v \equiv c_w \in \mathbb{R}$  on  $\Omega$  by (43) and, furthermore,  $v \equiv c_w$  on  $\Sigma_w$  by the trace theorem, i.e.,  $c_w = 0$ . Therefore,  $\|\cdot\|_{\Sigma_w}$  is a norm on  $H^1(\Omega)$  equivalent to  $\|\cdot\|_{H^1(\Omega)}$ , cf. [41, p. 153].

W.l.o.g., we assume  $\Psi_\xi(0) = 0$  for all  $\xi \in \Sigma_{\text{in}}$  and finish the proof by showing that there is a  $c_\delta > 0$  such that

$$c_\delta \|v\|_{L^2(\Sigma_w)}^2 \leq \int_{\Sigma_{\text{in}}} \Psi_\xi(v) \quad \forall v \in H^1(\Omega).$$

We prove that  $c_\delta v^2 \leq \Psi_\xi(v)$  for all  $\xi \in \Sigma_w$ . For all  $\xi \in \Sigma_{\text{in}}$ ,  $\Psi_\xi(0)$  is the minimum of  $\Psi_\xi$  since  $\Psi'_\xi(0) = \tau(\kappa_\xi^*)_\delta(0) = 0$  by (8) and (39). Therefore, it remains to show that  $2c_\delta v \leq \tau(\kappa_\xi^*)_\delta(v)$  for  $v \in \mathbb{R}$  and all  $\xi \in \Sigma_w$  or, equivalently, that  $\tau(\kappa_\xi^*)'_\delta \geq 2c_\delta$  for  $\xi \in \Sigma_w$ . By (8) and (39) we can choose  $c_\delta \leq \frac{\tau}{2c k_\delta(1)} \min\{1, \frac{w_0}{\sigma}\}$ .  $\square$

### 4.3 Finite element discretization and monotone multigrid

In the following, we shortly indicate our space discretization and numerical solution technique for (40). For more details we refer the reader to [10, Sec. 2.5, 3.4.5]. We discretize (40) or (41), respectively, by piecewise linear finite elements. For simplicity, we consider the two-dimensional case of a polygonal domain  $\Omega \subset \mathbb{R}^2$ . Let  $\mathcal{T}_j$ ,  $j \in \mathbb{N}_0$ , be a conforming triangulation of  $\Omega$ . The set of all vertices of the triangles in  $\mathcal{T}_j$  is denoted by  $\mathcal{N}_j$ , the set of those vertices lying on  $\bar{\Sigma}_{\text{in}}$  and  $\bar{\Sigma}_{\text{out}}$  shall be called  $\mathcal{N}_j^{\text{in}}$  and  $\mathcal{N}_j^{\text{out}}$ , respectively. The triangulation shall resolve the parts of the boundary corresponding to different boundary conditions.

Let  $\mathcal{S}_j \subset H^1(\Omega)$  be the finite element space of all continuous functions in  $H^1(\Omega)$  which are affine on each triangle  $t \in \mathcal{T}_j$ . To discretize the surface water level  $w$  we consider the trace grid  $\mathcal{T}_j^{\text{in}}$  of  $\mathcal{T}_j$  on  $\bar{\Sigma}_{\text{in}}$ , and its dual grid  $\mathcal{T}_j^{\text{in},*}$ . On  $\mathcal{T}_j^{\text{in},*}$  we define the finite element space  $\mathcal{S}_j^{\text{in}}$  as the space of all functions that are constant on each element. This is a suitable space for  $w$ , as the evolution equation (3) does not involve space derivatives. Note that each element of  $\mathcal{T}_j^{\text{in},*}$  corresponds uniquely to a vertex  $q$  of  $\mathcal{N}_j^{\text{in}}$  which we call element center. We will therefore sometimes use vertices of  $\mathcal{N}_j^{\text{in}}$  to label elements in  $\mathcal{T}_j^{\text{in},*}$ . The nodal basis function corresponding to a node  $q \in \mathcal{N}_j$  shall be called  $\lambda_q^{(j)}$ , and the (piecewise constant) basis function corresponding to an element  $q$  of  $\mathcal{T}_j^{\text{in},*}$  will be called  $\lambda_q^{\text{in},(j)}$ .

The finite dimensional analogue of  $\mathcal{K}$  is the convex set

$$\mathcal{K}_j := \{v \in \mathcal{S}_j : v(q) \geq \Phi(0) \quad \forall q \in \mathcal{N}_j \wedge v(q) \leq 0 \quad \forall q \in \mathcal{N}_j^{\text{out}}\}.$$

We define the  $\mathcal{S}_j$ -interpolation  $I_{\mathcal{S}_j} : C(\bar{\Omega}) \rightarrow \mathcal{S}_j$  by evaluation at the vertices, and the  $\mathcal{S}_j^{\text{in}}$ -interpolation  $I_{\mathcal{S}_j^{\text{in}}} : C(\bar{\Sigma}_{\text{in}}) \rightarrow \mathcal{S}_j^{\text{in}}$  by evaluation at the element centers of  $\mathcal{T}_j^{\text{in},*}$ . This allows to define the finite element discretization of  $F$  in (41) by

$$\begin{aligned} F_j : v &\mapsto \int_{\Omega} I_{\mathcal{S}_j} \Psi_x(v) + \int_{\Sigma_{\text{in}}} I_{\mathcal{S}_j^{\text{in}}} \Psi_{\xi}(v) + \frac{\tau}{2} \int_{\Omega} |\nabla v|^2 \\ &\quad - \int_{\Omega} \Phi^{-1}(u^n) \cdot v + \tau \int_{\Omega} k(\Phi^{-1}(u^n)) \mathbf{e} \cdot \nabla v - \tau \int_{\Sigma_{\text{in}}} \frac{w_j^n}{c} \cdot v \quad \forall v \in \mathcal{K}_j, \end{aligned} \quad (45)$$

where we assume that a discrete surface water level  $w_j^n \in \mathcal{S}_j^{\text{in}}$  is available and also used in the definition of  $(\kappa_{\xi}^*)_{\delta}$  and, equivalently, of  $\Psi_{\xi}$ . The discrete finite element solution  $u_j^n \in \mathcal{S}_j$  is assumed to be known, too. Then, with the weights

$$h_q := \int_{\Omega} \lambda_q^{(j)} \quad \text{and} \quad h_q^{\text{in}} := \int_{\Sigma_{\text{in}}} \lambda_q^{\text{in},(j)},$$

we can note the following.

**Theorem 4.2.** *Assume that the conditions of Theorem 4.1 are satisfied. Then the discrete variational inequality for  $u_j^{n+1} \in \mathcal{K}_j$  :*

$$\begin{aligned} &\sum_{q \in \mathcal{N}_j} \left[ \Phi^{-1}(u_j^{n+1}(q)) - \tau f(\Phi^{-1}(u_j^{n+1}(q))) \right] (v(q) - u_j^{n+1}(q)) h_q \\ &\quad + \tau \sum_{q \in \mathcal{N}_j^{\text{in}}} (\kappa_{\xi}^*)_{\delta}(u_j^{n+1}(q)) (v(q) - u_j^{n+1}(q)) h_q^{\text{in}} \\ &\quad + \tau \int_{\Omega} \nabla u_j^{n+1} \cdot \nabla (v - u_j^{n+1}) - \int_{\Omega} \Phi^{-1}(u_j^n) \cdot (v - u_j^{n+1}) \\ &\quad + \tau \int_{\Omega} k(\Phi^{-1}(u_j^n)) \mathbf{e} \cdot \nabla (v - u_j^{n+1}) - \tau \int_{\Sigma_{\text{in}}} \frac{w_j^n}{c} \cdot (v - u_j^{n+1}) \geq 0 \quad \forall v \in \mathcal{K}_j \end{aligned} \quad (46)$$

is equivalent to the discrete convex minimization problem

$$u_j^{n+1} \in \mathcal{K}_j : \quad F_j(u_j^{n+1}) \leq F_j(v) \quad \forall v \in \mathcal{K}_j. \quad (47)$$

If  $w^n(q) > 0$  holds for a  $q \in \mathcal{N}_j^{\text{in}}$ , then (47) is uniquely solvable.

The proof of this theorem for  $\mathcal{K}_j \subset \mathcal{S}_j$  is analogous to the one of its continuous counterpart, Theorem 4.1, for  $\mathcal{K} \subset H^1(\Omega)$ . For this we refer again to [10, Sec. 2.5] and [12]. We only point out that the second integral in the definition (45) of  $F_j$  is the same as if  $\mathcal{S}_j^{\text{in}}$  is replaced by the space of piecewise linear finite elements on  $\mathcal{T}_j^{\text{in}}$ . As for (40), well-posedness can also be established for (46). Under reasonable assumptions, one also obtains a convergence result for the discretization.

**Theorem 4.3.** *Let the conditions of Theorem 4.1 be satisfied and  $\mathcal{T}_j$ ,  $j \geq 0$ , be shape regular and  $h_j = \max_{t \in \mathcal{T}_j} \text{diam } t \rightarrow 0$  for  $j \rightarrow \infty$ . Let  $w_j^n \in L^2(\Sigma_{\text{in}})$ ,  $j \geq 0$ , be given with  $w_j^n \rightarrow w^n$  in  $L^2(\Sigma_{\text{in}})$  for  $j \rightarrow \infty$ . Furthermore, we assume  $C^\infty(\overline{\Omega}) \cap \mathcal{K}$  to be dense in  $\mathcal{K}$ . We define  $s^{n+1} = \Phi^{-1}(u^{n+1})$  and, with  $\kappa_\delta$  obtained by regularization of  $\kappa$  according Assumption 3.2, we define  $p^{n+1} = \kappa_\delta^{-1}(u^{n+1})$  (analogously,  $s_j^{n+1}$  and  $p_j^{n+1}$  are defined). Then, for  $j \rightarrow \infty$ , we have the convergence*

$$u_j^{n+1} \rightarrow u^{n+1} \quad \text{in } H^1(\Omega), \quad (48)$$

$$p_j^{n+1} \rightarrow p^{n+1} \quad \text{in } H^1(\Omega) \quad \text{and} \quad I_{\mathcal{S}_j} p_j^{n+1} \rightarrow p^{n+1} \quad \text{in } L^2(\Omega). \quad (49)$$

In addition, we have

$$s_j^{n+1} \rightarrow s^{n+1} \quad \text{in } H^1(\Omega) \quad \text{and} \quad I_{\mathcal{S}_j} s_j^{n+1} \rightarrow s^{n+1} \quad \text{in } L^2(\Omega) \quad (50)$$

if  $\Phi^{-1} : (\Phi(0), \infty) \rightarrow \mathbb{R}$  is Lipschitz continuous for the first and Hölder continuous for the second statement, respectively.

The proof of (48) is an adaptation of the one given in [10, Thm. 2.5.9], where we need to use the norm defined in (44) instead of the  $H^1$ -seminorm as an equivalent norm in the solution space, which is  $H^1(\Omega)$  here. For the proofs of (49) and (50) see [13].

In terms of finite elements, a straightforward discretization of (38)<sub>7</sub> reads

$$w_j^{n+1} = \tau \left( r_j^{n+1} - I_{\mathcal{S}_j^{\text{in}}} g \left( u_{j|\mathcal{S}_j^{\text{in}}}^{n+1}, w_j^n \right) \right) + w_j^n \quad \text{in } \mathcal{S}_j^{\text{in}}.$$

Here, we can assume that  $u_j^{n+1}$  has been obtained by (46) and that the iterates  $r_j^{n+1}$ ,  $w_j^n \in \mathcal{S}_j^{\text{in}}$  are given.

Quadrature of the second integral in (46) is given by  $\mathcal{S}_j$ -interpolation of  $\Phi^{-1}(u_j^n)$ . In addition, with regard to the numerical treatment of the discretized problem (46), it is necessary to use an upwind technique for the gravitational (convective) term  $k(\Phi^{-1}(u^n))\mathbf{e}$  in the third integral. In the finite element context, this can be achieved by adding an artificial viscosity term to the discretized convection, cf. [10, Sec. 4.2] and [13].

The discrete convex obstacle problem arising from (46) is defined on a set  $\mathcal{K}_j$  which is a product of intervals. To such problems, monotone multigrid methods can be applied [28]. In short, these methods first minimize  $F_j$  by successive

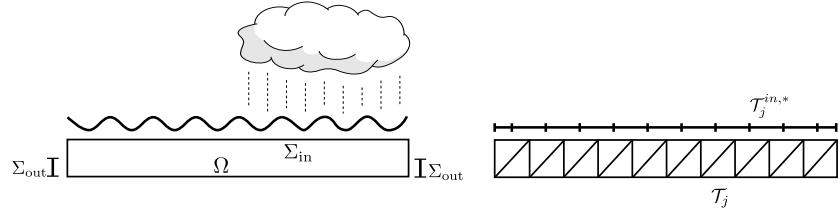


Figure 3: Left: Setting of the numerical example. Right: Coarse grids. The actual grids were created from these grids by four steps of uniform refinement.

one-dimensional minimization along the  $\lambda_q^{(j)}$ -directions for  $q \in \mathcal{N}_j$ . After application of this nonlinear Gauss–Seidel fine grid smoother, a suitable coarse grid correction given by constraint Newton linearization in the smooth regimes of the nonlinear system given by (46) accelerates convergence. Linear multigrid convergence rates are obtained asymptotically. For details concerning the application of this method to a Signorini-type and Robin problem like (46) we refer to [10, Sec. 2.7, 3.4.5].

In the next section, we will apply monotone multigrid to an example with Brooks–Corey parameter functions which do neither satisfy the conditions for the existence theorem 3.1 nor for the theorems given in this section which require Kirchhoff transformations given by regularized parameter functions. However, even though the equivalence result of Theorem 4.1 does not hold if  $(\kappa_\xi^*)_\delta$  is replaced by  $\kappa_\xi^*$  arising from the Brooks–Corey parametrization (since (40) is not well-defined), it turns out that (42) is still a well-posed convex minimization problem in this case. The same is true for the discretized problem (since  $\Phi(0) < u_j^{n+1}$ , one can even prove Theorem 4.2 here) so that (47) is well-posed and can be solved by monotone multigrid methods [10, Sec. 3.4.5].

## 5 Numerical example

We close the article with a numerical example involving the coupling of surface water to ground water, and a Signorini-type outflow condition. We conclude that our model and solution algorithm is well-suited for practical applications.

The problem setting is depicted schematically in the left part of Fig. 3. As the ground water domain, we use a  $10 \text{ m} \times 1 \text{ m}$  rectangle whose lower left corner shall be situated in the origin of  $\mathbb{R}^2$ . Zero-flow boundary conditions are prescribed on the left, right, and bottom sides, with the exception of two  $0.5 \text{ m}$  stretches  $\Sigma_{\text{out}}$  on the lower left and right, where we set the Signorini outflow conditions (2). The entire upper side  $\Sigma_{\text{in}}$  couples with the surface water model through the flux conditions (3) and (5). The resistance  $c$  is set to  $10^5 \text{ s}$  (cf. [42]), and for the threshold parameter  $\sigma$  we choose  $0.02 \text{ m}$ . We use sand as the porous medium, with soil parameters  $n = 0.437$  (porosity),  $s_m = 0.0458$  (residual saturation),  $s_M = 1$  (maximal saturation),  $p_b = -712.2 \text{ Pa}$  (bubbling pressure),  $\lambda = 0.694$  (pore size distribution factor), and  $K = 6.66 \cdot 10^{-9} \text{ m}^2$  (absolute permeability), taken from [36, Tables 5.3.2 and 5.5.5]. We set  $\mu = 1.002 \cdot 10^{-3} \text{ Pa s}$  for the dynamic viscosity of water. The parameters  $p_b$  and  $\lambda$  characterize the parameter functions  $p \mapsto s = p_c^{-1}(p)$  and  $s \mapsto k(s)$ , according to Brooks–Corey [16, 40]

and Burdine [17] of the form

$$s = p_c^{-1}(p) = \begin{cases} s_m + (s_M - s_m) \left( \frac{p}{p_b} \right)^{-\lambda} & \text{for } p \leq p_b, \\ s_M & \text{for } p \geq p_b, \end{cases} \quad (51)$$

and

$$k(s) = K \left( \frac{s - s_m}{s_M - s_m} \right)^{3+\frac{2}{\lambda}}, \quad s \in [s_m, s_M]. \quad (52)$$

The domain is discretized with a uniform triangle grid on  $161 \times 17$  vertices, giving a mesh size of  $h = \sqrt{2}/16$  m. We construct this grid by four uniform refinement steps of a uniform  $10 \times 1$  grid. As a by-product we obtain a grid hierarchy suitable for our multigrid solver. To discretize the surface water model, we use the dual grid  $\mathcal{T}_j^{\text{in},*}$  of the trace of the subsurface grid on  $\Sigma_{\text{in}}$ . The grids are depicted schematically on the right of Fig. 3. For the implementation we used the DUNE libraries [6] together with the domain decomposition module `dune-grid-glue` [7].

We start with the initial pressure  $-2 \cdot 10^4$  Pa in the domain, which, taking the material parameters and (51) into account, corresponds to the constant initial saturation 0.1401. The initial surface water height is set to a constant 0 m, and we prescribe a constant rainfall of  $30 \text{ mm/h} \approx 8.33 \cdot 10^{-6} \text{ m/s}$  on  $[5, 10] \times \{1\}$ , and zero elsewhere. The evolution is simulated up to a final time  $T = 350\,000$  s, which equals roughly four days. For the time step size  $\tau$  we choose 100 s. This is well below the upper bound on the time step size implied by the CFL condition

$$\tau < h \frac{n\mu}{\varrho g \sup_{s \in [0,1]} |k'(s)|} = h \frac{n\mu}{K \varrho g (3 + 2/\lambda)},$$

cf. [10, p. 223], which in our setting evaluates to approximately  $\tau < h \cdot 1.14 \cdot 10^3 \text{ s/m} = 200$  s. For each time step we solve the spatial problem (46) to machine precision by the monotone multigrid method described in [13]. We omit the solver convergence rates and refer the interested reader to [13] for robustness and efficiency studies. While solving to machine precision usually means more work than necessary, it eliminates the effect of solver inexactness from the following observations.

Fig. 4 shows a few snapshots from the evolution. Initially, surface water accumulates on the right, where there is rainfall, and infiltrates into the soil. A saturation front traverses the domain. When the right part of  $\Omega$  is almost filled, a hydrostatic pressure distribution starts to build up there, and the water starts to infiltrate the left part of  $\Omega$ . Around time step 2429, the ground is fully saturated. Most of the water entering the domain leaves it through  $\Sigma_{\text{out}}$ , while a small part exfiltrates on the left part of  $\Sigma_{\text{in}}$ , and leads to a rising surface water level there. At the last time step, the subsurface pressure field is contained in the interval  $[0 \text{ Pa}, 2.53 \cdot 10^4 \text{ Pa}]$ , and the surface water height is contained in  $[0.23 \text{ m}, 1.89 \text{ m}]$ .

In a second step, we try to numerically assess the time step bound (21). The proof of Theorem 3.11 shows that a nonnegative surface water height  $w^n$  at time step  $n$  implies a nonnegative surface water height  $w^{n+1}$  at time step  $t_{n+1}$  if the time step size  $\tau = t_{n+1} - t_n$  is bounded by (21), which states

$$\tau \leq \min\{c, \theta_1(x), \theta_2(x)\} \quad \text{for almost all } x \in \Sigma_{\text{in}}, \quad (53)$$

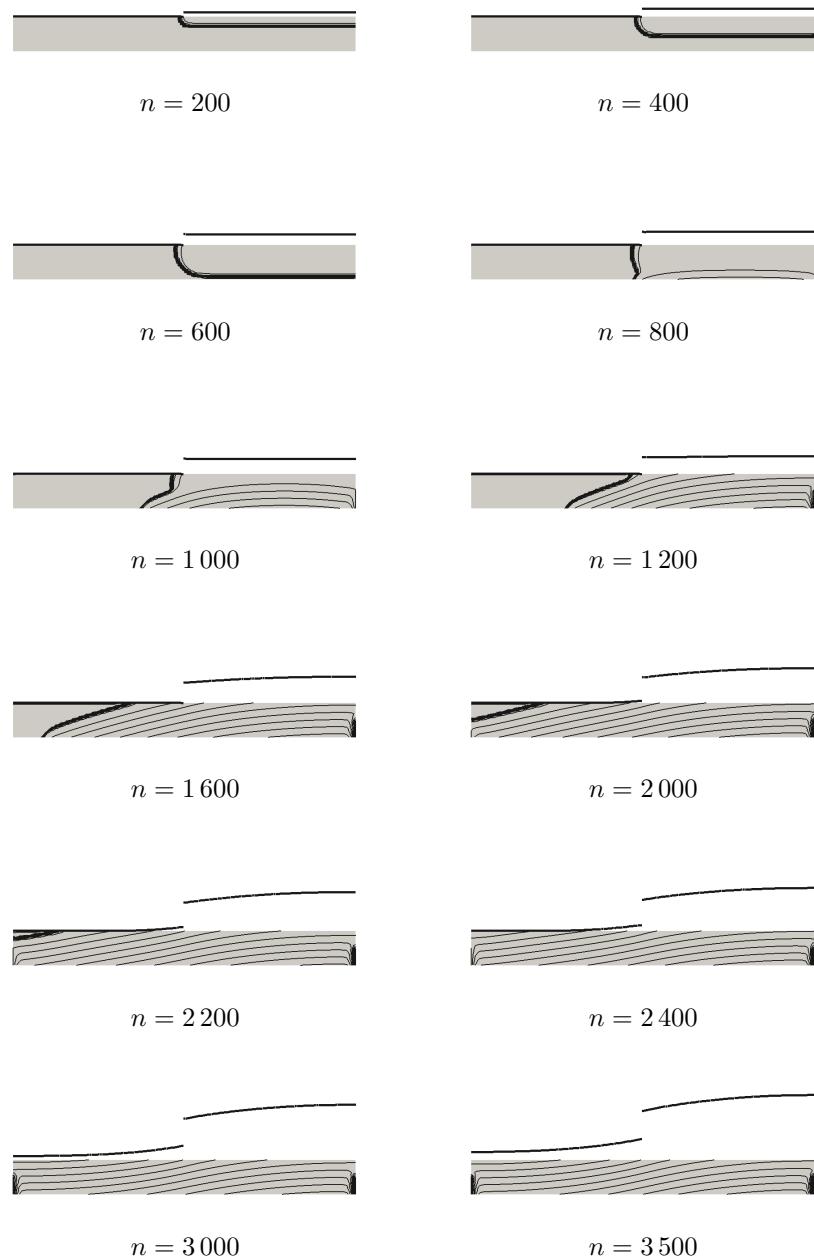


Figure 4: Snapshots of the numerical evolution at time steps  $t_n = n \cdot \tau = n \cdot 100 \text{ s}$  as indicated, with the top black line giving the surface water height, and isolines of the pressure  $p$  at integral multiples of 2 kPa.

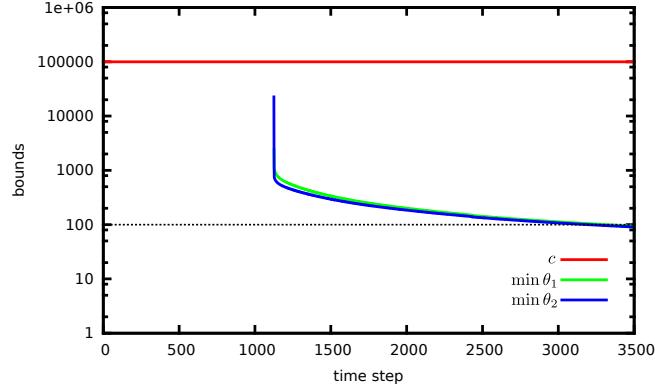


Figure 5: The three terms of the time step restriction, plotted as functions over the time step number; dotted line: used time step size  $\tau = 100$  s.

where

$$\theta_1(x) = \begin{cases} \frac{c\sigma}{\sigma - cr(x, t_n) + h(u^n(x))} & \text{if this term is nonnegative,} \\ \infty & \text{else} \end{cases}$$

and

$$\theta_2(x) = \begin{cases} \frac{c\sigma}{\sigma + h(u^n(x))} & \text{if this term is nonnegative,} \\ \infty & \text{else.} \end{cases}$$

Note that by construction, the minimum in (53) is always a finite nonnegative number.

In Fig. 5 we have plotted  $c$ ,  $\min_{x \in \Sigma_{\text{in}}} \theta_1(x)$  and  $\min_{x \in \Sigma_{\text{in}}} \theta_2(x)$ , with  $u^n$  replaced by  $u_j^n$  for  $j = 4$ , as functions of the time step number  $n$ . We observe that our time step is below the bound (21) for almost all time steps. In accordance with the continuous theory we observe nonnegativity of the solution everywhere, with one notable exception. A surface water element with initial value  $w_p = 0$ , but next to an element  $q$  with  $w_q > 0$  will drop below zero if the subsurface pressure is negative there. This is a discretization effect. Unlike in the continuous case, the surface water levels  $w_p$  and  $w_q$  couple through a subsurface basis function. Since  $w_q > 0$  and the subsurface pressure is negative, there will be a flux into the subsurface which will in turn suck water out of the element  $p$ , making  $w_p$  negative. Through this effect, negative surface water heights down to  $-0.013$  m were produced in our example.

It is plausible to try to use the bound (53), with  $u^n$  replaced by  $u_j^n$ , as a time step control mechanism. To check whether this bound is sharp, we have recomputed our example with the time step sizes  $\tau = 50$  s, 100 s, 200 s, 400 s, 800 s, 1600 s, 3200 s. Only at the last value of  $\tau$  we do observe instabilities presumably caused by violations of the CFL condition. As it turns out, the graphs in Fig. 5 are virtually independent of the time step size. Also, we do not see negative water levels (except for the discretization effect described above) for any of the time step sizes considered. Hence we conclude that the sufficient

condition (53) in the spatially continuous case is generally too tight when applied to fully discretized problems.

## Acknowledgments

We thank Ben Schweizer for fruitful discussions during this work.

## References

- [1] H.W. Alt and S. Luckhaus. Quasilinear elliptic–parabolic differential equations. *Math. Z.*, 183:311–341, 1983.
- [2] H.W. Alt, S. Luckhaus, and A. Visintin. On nonstationary flow through porous media. *Ann. Math. Pura Appl.*, 136:303–316, 1984.
- [3] T. Arbogast, M. Wheeler, and N.-Y. Zhang. A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media. *SIAM J. Numer. Anal.*, 33:1669–1687, 1996.
- [4] C. Bardos, A. Y. le Roux, and J.-C. Nédélec. First order quasilinear equations with boundary conditions. *Comm. Partial Differential Equations*, 4 (9):1017–1034, 1979.
- [5] P. Bastian, O. Ippisch, F. Rezanezhad, H.J. Vogel, and K. Roth. Numerical simulation and experimental studies of unsaturated water flow in heterogeneous systems. In W. Jäger, R. Rannacher, and J. Warnatz, editors, *Reactive Flows, Diffusion and Transport*, pages 579–598. Springer, 2005.
- [6] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic interface for parallel and adaptive scientific computing. Part II: Implementation and tests in DUNE. *Computing*, 82(2–3):121–138, 2008.
- [7] P. Bastian, G. Buse, and O. Sander. Infrastructure for the coupling of Dune grids. In *Proc. of ENUMATH 2009*, pages 107–114. Springer, 2010.
- [8] P. Bastian, H. Berninger, A. Dedner, C. Engwer, P. Henning, R. Kornhuber, D. Kröner, M. Ohlberger, O. Sander, G. Schiffler, N. Shokina, and K. Smetana. Adaptive modelling of coupled hydrological processes with application in water management. In M. Günther et al., editor, *Progress in Industrial Mathematics at ECMI 2010*, volume 17 of *Mathematics in Industry*, pages 559–565. Springer, 2012.
- [9] J. Bear. *Dynamics of fluids in porous media*, volume 2. Dover Publications, 1988.
- [10] H. Berninger. *Domain Decomposition Methods for Elliptic Problems with Jumping Nonlinearities and Application to the Richards Equation*. PhD thesis, Freie Universität Berlin, 2007.
- [11] H. Berninger. Non-overlapping domain decomposition for the Richards equation via superposition operators. In M. Bercovier, M.J. Gander, R. Kornhuber, and O. Widlund, editors, *Domain Decomposition Methods in*

*Science and Engineering XVIII*, volume 70 of *LNCSE*, pages 169–176. Springer, 2009.

- [12] H. Berninger and O. Sander. Substructuring of a Signorini-type problem and Robin’s method for the Richards equation in heterogeneous soil. *Comput. Vis. Sci.*, 13(5):187–205, 2010.
- [13] H. Berninger, R. Kornhuber, and O. Sander. Fast and robust numerical solution of the Richards equation in homogeneous soil. *SIAM J. Numer. Anal.*, 49(6):2576–2597, 2011.
- [14] H. Berninger, R. Kornhuber, and O. Sander. Heteogeneous substructuring methods for coupled surface and subsurface flow. In *Proceedings of the 20th conference on domain decomposition methods*, accepted.
- [15] F. Brezzi and G. Gilardi. Functional spaces. In H. Kardestuncer and D.H. Norrie, editors, *Finite Element Handbook*, chapter 2 (part 1), pages 1.29–1.75. Springer, 1987.
- [16] R.J. Brooks and A.T. Corey. Hydraulic properties of porous media. Technical Report Hydrology Paper No. 3, Colorado State University, Civil Engineering Department, Fort Collins, 1964.
- [17] N.T. Burdine. Relative permeability calculations from pore-size distribution data. *Petr. Trans., Am. Inst. Mining Metall. Eng.*, 198:71–77, 1953.
- [18] X. Chen, A. Friedman, and T. Kimura. Nonstationary filtration in partially saturated porous media. *European J. Appl. Math.*, 5(3):405–429, 1994.
- [19] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 2, Functional and Variational Methods. Springer, 2000.
- [20] C. Dawson. Analysis of discontinuous finite element methods for ground water/surface water coupling. *SIAM J. Numer. Anal.*, 44(4):1375–1404, 2006.
- [21] C. Dawson. A continuous/discontinuous Galerkin framework for modeling coupled subsurface and surface water flow. *Comput. Geosci.*, 12:451–472, 2008.
- [22] R. Eymard, M. Gutnic, and D. Hilhorst. The finite volume method for Richards equation. *Comput. Geosci.*, 3(3–4):259–294, 1999.
- [23] J. Filo. Finite time of stabilization in the one-dimensional problem of non-steady filtration. *Math. Methods Appl. Sci.*, 19(7):529–554, 1996.
- [24] J. Filo and S. Luckhaus. Modelling surface runoff and infiltration of rain by an elliptic-parabolic equation coupled with a first-order equation on the boundary. *Arch. Ration. Mech. Anal.*, 146(2):157–182, 1999.
- [25] P.A. Forsyth and M.C. Kropinski. Monotonicity considerations for saturated-unsaturated subsurface flow. *SIAM J. Sci. Comput.*, 18(5):1328–1354, 1997.

- [26] J. Fuhrmann. *Zur Verwendung von Mehrgitterverfahren bei der numerischen Behandlung elliptischer partieller Differentialgleichungen mit variablen Koeffizienten*. PhD thesis, TU Chemnitz–Zwickau, 1994.
- [27] J. Fuhrmann and H. Langmach. Stability and existence of solutions of time-implicit finite volume schemes for viscous nonlinear conservation laws. *Appl. Numer. Math.*, 37(1–2):201–230, 2001.
- [28] R. Kornhuber. On constrained Newton linearization and multigrid for variational inequalities. *Numer. Math.*, 91:699–721, 2002.
- [29] S.N. Kružkov. First order quasilinear equations in several independent variables. *Sb. Math.*, 10:217–243, 1970.
- [30] O.A. Ladyzenskaja, V.A. Solonnikov, and N.N. Ural'ceva. *Linear and quasi-linear equations of parabolic type*. Amer. Math. Soc., 1968.
- [31] M. Ohlberger and B. Schweizer. Modelling of interfaces in unsaturated porous media. *Discrete Contin. Dyn. Syst.*, (Dynamical Systems and Differential Equations. Proceedings of the 6th AIMS International Conference, suppl.):794–803, 2007.
- [32] F. Otto.  $L^1$ -contraction and uniqueness for quasilinear elliptic–parabolic equations. *J. Differ. Equations*, 131(1):20–38, 1996.
- [33] F. Otto.  $L^1$ -contraction and uniqueness for unstationary saturated–unsaturated porous media flow. *Adv. Math. Sci. Appl.*, 7(2):537–553, 1997.
- [34] I.S. Pop and B. Schweizer. Regularization schemes for degenerate Richards equations and outflow conditions. *Math. Models Methods Appl.*, 21(8):1685–1712, 2011.
- [35] F. Radu, I.S. Pop, and P. Knabner. Order of convergence estimates for an Euler implicit, mixed finite element discretization of Richards' equation. *SIAM J. Numer. Anal.*, 42(4):1452–1478, 2004.
- [36] W.J. Rawls, L.R. Ahuja, D.L. Brakensiek, and A. Shirmohammadi. Infiltration and soil water movement. In D.R. Maidment, editor, *Handbook of Hydrology*, chapter 5. McGraw–Hill, 1993.
- [37] B. Schweizer. Regularization of outflow problems in unsaturated porous media with dry regions. *J. Differential Equations*, 237(2):278–306, 2007.
- [38] K. Smetana. Betrachtung von Infiltrationsprozessen und der Einphasenströmung im ungesättigten porösen Medium. Modellierung mit Hilfe der Richards-Gleichung gekoppelt mit einer gewöhnlichen Differentialgleichung und Beweis der Existenz von Lösungen. Master's thesis, Albert-Ludwigs-Universität Freiburg, 2008.
- [39] P. Sochala, A. Ern, and S. Piperno. Mass conservative BDF-discontinuous Galerkin/explicit finite volume schemes for coupling subsurface and overland flows. *Comput. Methods Appl. Mech. Engrg.*, 198:2122–2136, 2009.
- [40] M.T. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 44:892–898, 1980.

- [41] D. Werner. *Funktionalanalysis*. Springer, 5th edition, 2005.
- [42] B. Wiese and G. Nützmann. Transient leakance and infiltration characteristics during lake bank filtration. *Ground Water*, 47(1):57–68, 2009.